

Agentes de IA na empresa

O que delegar primeiro, como desenhar agentes com uma função clara e como organizar equipes de agentes com supervisão humana.



D.B. Conhecimento

Coleção Inteligência Aplicada · Volume 05

■ AGENTES DE IA NA EMPRESA: DELEGAR, ORQUESTRAR E SUPERVISIONAR

O que delegar primeiro, como desenhar agentes com uma função clara e como organizar equipes de agentes com supervisão humana.

1ª edição · outubro de 2026 · Autoria: D.B. Soluções Digitais.

■ USO E DISTRIBUIÇÃO

Material educativo de distribuição gratuita. É permitido ler, imprimir e compartilhar este arquivo sem alterações, inclusive dentro de empresas. A venda é proibida. Trechos podem ser citados com indicação da fonte: "D.B. Soluções Digitais, D.B. Conhecimento".

■ AVISO

O conteúdo tem finalidade educativa e não substitui consultoria jurídica, contábil, tributária ou de segurança da informação. Estudos de caso e exemplos numéricos são ilustrativos, com empresas e valores fictícios. Marcas e produtos citados pertencem aos respectivos titulares; a menção não indica parceria nem recomendação comercial. Ferramentas de IA mudam rapidamente: confirme recursos, preços e termos nas fontes oficiais.

■ CRÉDITOS

Texto, diagramas e projeto gráfico: D.B. Soluções Digitais. Imagens geradas para este projeto com inteligência artificial. Tipografia: Satoshi (Indian Type Foundry, licença ITF FFL), Source Serif 4, Manrope, Instrument Serif e JetBrains Mono (SIL Open Font License).

© 2026 D.B. Soluções Digitais. Além do código.

Delegar é uma competência

Agentes de inteligência artificial deixaram de ser tema de laboratório. Fornecedores de modelos, plataformas de automação e empresas de software lançam, quase toda semana, agentes que prometem pesquisar clientes, responder e-mails, montar propostas, publicar conteúdo e até trabalhar enquanto a equipe dorme. Algumas dessas promessas já funcionam no dia a dia de empresas reais. Outras ainda dependem de condições que as demonstrações não mostram.

A distância entre a demonstração e a operação está quase sempre nos mesmos pontos: tarefas mal definidas, ferramentas com permissões amplas demais, ausência de critérios de qualidade, falta de supervisão e nenhuma forma de medir se o agente está ajudando ou criando trabalho extra. Esta apostila trata justamente desses pontos.

Você vai entender o que é um agente e como ele difere de um assistente e de uma automação, como escolher as primeiras tarefas a delegar, como escrever o briefing de um agente, quais padrões de orquestração usar quando vários agentes trabalham juntos, como supervisionar e avaliar, quais riscos de segurança merecem atenção e como implantar com indicadores.

Ao longo do texto, aparecem referências a organizações de agentes por departamento, com dezenas de agentes especializados em vendas, marketing e operações, um formato que ganhou visibilidade em bibliotecas abertas sobre IA aplicada a negócios. Elas são úteis como mapa de possibilidades. Mas a recomendação desta apostila é clara: comece com um agente, uma tarefa e uma pessoa responsável. Equipes de agentes são o resultado de várias pequenas implantações bem-sucedidas, não o ponto de partida.

O que você vai encontrar

—	Apresentação	3
01	O problema: a promessa e a prática	7
	O primeiro dia de um agente	7
	Por que as delegações falham	8
	O custo do erro como critério central	8
02	O que é um agente de IA	10
	Definição prática	10
	Componentes de um agente	11
	Por dentro de uma execução	11
	Níveis de autonomia	12
	Agente, assistente e automação	13
03	O que delegar primeiro	14
	A régua de duas perguntas	14
	Dez tarefas que costumam funcionar bem	15
04	O briefing de um agente	18
	As partes do briefing	18
	Um exemplo completo	19
	As instruções de sistema, comentadas	19
	Ferramentas bem desenhadas	20
	Uma função por agente	21
05	Equipes de agentes: padrões de orquestração	22
	Quando um agente não basta	22
	Erros comuns em equipes de agentes	25
06	Ferramentas e plataformas	26
	Categorias de ferramentas	26
	Critérios de escolha	27
	Estimando o custo de um agente	27

07	Supervisão, avaliação e segurança	29
	Registos de execução	29
	Avaliação antes e depois de publicar	29
	Riscos de segurança	31
	Regras aprendidas da forma difícil	32
	Quando algo dá errado: resposta a incidentes	32
08	Implantação e indicadores	34
	O ritual semanal de revisão	35
09	Estudo de caso ilustrativo e exercícios	37
10	Como a D.B. pode atuar e conclusão	39
	Conclusão	39
—	Glossário	40
—	Referências	41

Ao final desta apostila, você será capaz de:

- 01 Definir agente de IA e diferenciá-lo de automações e assistentes.
- 02 Selecionar tarefas para delegação usando valor e custo do erro.
- 03 Escrever o briefing completo de um agente.
- 04 Escolher padrões de orquestração para equipes de agentes.
- 05 Implantar supervisão, aprovações, registros e avaliação contínua.
- 06 Reconhecer e mitigar os principais riscos de segurança de agentes.

Para quem é este material. gestores que querem delegar tarefas repetitivas; equipes de vendas, marketing, atendimento e operações; profissionais de tecnologia que vão construir agentes.

Como aproveitar melhor. Leia um capítulo por vez, faça os exercícios com um processo real da sua empresa e use os checklists como roteiro de reunião. Os estudos de caso são ilustrativos: servem para mostrar o raciocínio, não para prometer resultados.

O problema: a promessa e a prática

Delegar a um agente é delegar a um colega novo, muito rápido e sem bom senso próprio.

■ O primeiro dia de um agente

O roteiro de fracasso é conhecido. A empresa contrata uma plataforma de agentes, cria um agente com nome simpático e uma instrução ampla, como “cuide da minha caixa de e-mail” ou “faça a prospecção de clientes”. Conecta o agente ao e-mail, ao CRM e à agenda. No primeiro dia, ele arquiva mensagens importantes, responde a um fornecedor com informação errada ou envia mensagens genéricas a uma lista de contatos. A equipe perde a confiança, desliga o agente e conclui que “agentes ainda não funcionam”.

O problema raramente é o modelo. É a delegação. Uma pessoa recém-contratada que recebesse a mesma instrução, com as mesmas permissões e sem treinamento, também causaria estragos. A diferença é que o agente causa estragos mais rápido e sem perceber o constrangimento.

■ Por que as delegações falham

- **Tarefa vaga.** “Cuidar do e-mail” abrange centenas de decisões diferentes, de excluir spam a responder a uma reclamação jurídica.
- **Sem critério de pronto.** O agente não sabe quando parar nem como verificar se acertou.
- **Permissões amplas.** O agente pode enviar, apagar e alterar, quando bastaria redigir e sugerir.
- **Sem contexto.** O agente não conhece a empresa, os clientes, a voz e as regras.
- **Sem escalonamento.** Não há regra sobre quando pedir ajuda a uma pessoa.
- **Sem medição.** Ninguém compara o trabalho do agente com uma linha de base.

● COMPARATIVO	
DELEGAÇÃO RUIM	DELEGAÇÃO BEM FEITA
Cuide da caixa de entrada	→ Classifique as mensagens novas em cinco categorias e rascunhe respostas para dúvidas frequentes
Acesso total ao e-mail	→ Leitura da caixa e escrita apenas em rascunhos
Sem regras de exceção	→ Reclamações e temas jurídicos vão para a coordenadora
Sem verificação	→ Cada rascunho cita a política usada como base
Sucesso medido pela impressão	→ Taxa de rascunhos aprovados sem alteração, por semana

Figura 1.1 — Delegação ruim e delegação bem feita

■ O custo do erro como critério central

Em qualquer decisão sobre agentes, a pergunta mais importante não é “ele consegue fazer?”, mas “o que acontece quando ele erra?”. Um erro em um resumo interno custa minutos. Um erro em uma cobrança, em uma

mensagem a um cliente importante ou em um registro contábil pode custar dinheiro, reputação e tempo de correção. A estratégia de adoção deve começar pelas tarefas em que errar é barato e corrigir é fácil, e avançar à medida que a confiabilidade é demonstrada com dados.

**EM RESUMO**

Agentes falham principalmente por delegação mal feita: tarefas vagas, permissões amplas, falta de contexto, de critérios, de escalonamento e de medição. O custo de um erro deve orientar o que delegar primeiro e quanto controle manter.

O que é um agente de IA

Um sistema que persegue um objetivo, decide os próximos passos e usa ferramentas, dentro de limites.

■ Definição prática

Na literatura de inteligência artificial, agente é qualquer sistema que percebe o ambiente e age sobre ele para atingir objetivos, como definem Stuart Russell e Peter Norvig em seu livro de referência. No contexto atual, o termo se refere a sistemas construídos com modelos de linguagem que recebem um objetivo, decidem quais passos seguir, usam ferramentas para obter informações e executar ações, observam os resultados e continuam até concluir ou escalar.

A Anthropic, em um texto amplamente citado sobre a construção de agentes eficazes, recomenda distinguir **fluxos de trabalho**, em que modelos e ferramentas seguem caminhos definidos em código, de **agentes**, em que o próprio modelo dirige o processo. A recomendação prática do texto é buscar a solução mais simples possível e só aumentar a complexidade quando ela trazer ganho comprovado.

■ Componentes de um agente



Figura 2.1 — Componentes de um agente de IA

Cada componente é uma decisão de projeto. As ferramentas definem o que o agente pode fazer no mundo. As instruções definem como ele deve fazer. O contexto define o que ele sabe. A supervisão define quando uma pessoa entra.

■ Por dentro de uma execução

Para entender como um agente trabalha, vale acompanhar uma execução passo a passo. O exemplo abaixo mostra o rastro simplificado de um agente que prepara o resumo de pendências de um cliente antes de uma reunião.

Tabela 2.1 – Rastro simplificado de uma execução de agente (exemplo ilustrativo)

Passo	O agente pensa	Ação	Resultado observado
1	Preciso saber quem é o cliente e a data da reunião	Lê o evento da agenda	Cliente identificado, reunião amanhã às 10h
2	Quais pedidos estão abertos?	Consulta o sistema de gestão (somente leitura)	Três pedidos, um com entrega atrasada
3	Há reclamações recentes?	Busca chamados dos últimos 60 dias	Um chamado aberto sobre o pedido atrasado
4	Há faturas vencidas?	Consulta o financeiro (somente leitura)	Nenhuma pendência financeira
5	Tenho o suficiente para o resumo	Escreve o resumo no modelo, citando cada fonte	Arquivo salvo na pasta de saídas
6	O resumo atende aos critérios?	Confere se cada fato tem fonte e se não há dados de outros clientes	Verificação aprovada; tarefa concluída

Esse rastro revela três características importantes. O agente decide a ordem das consultas, em vez de seguir um roteiro fixo. Cada ação usa uma ferramenta com permissão limitada. E a execução termina com uma verificação explícita, prevista nas instruções. Se o passo 3 tivesse encontrado uma reclamação jurídica, uma regra de escalonamento faria o agente parar e avisar o responsável.

■ **Níveis de autonomia**

Pesquisadores de interação entre pessoas e automação, como Raja Parasuraman, Thomas Sheridan e Christopher Wickens, propuseram há mais de duas décadas que a automação não é uma questão de “tudo ou nada”, mas de níveis. A ideia se aplica perfeitamente aos agentes atuais. Uma escala simples ajuda a decidir quanto controle manter:

Tabela 2.2 – Níveis de autonomia para agentes de IA

Nível	O agente	A pessoa	Exemplo
1. Sugere	Analisa e propõe	Decide e executa	Sugere a categoria de cada chamado
2. Prepara	Executa até o rascunho	Revisa e aprova cada caso	Redige respostas que ficam em rascunho
3. Executa com amostragem	Executa casos padrão	Revisa uma amostra e as exceções	Responde dúvidas frequentes, com auditoria diária
4. Executa com limites	Executa dentro de regras e orçamentos	Acompanha indicadores e alertas	Reprocessa pedidos rejeitados por dados incompletos
5. Autônomo	Define e executa ações amplas	Intervém apenas em incidentes	Raramente adequado em processos com clientes ou dinheiro

A maior parte das tarefas empresariais com valor real fica entre os níveis 2 e 4. Uma tarefa sobe de nível apenas quando os indicadores mostram que a qualidade se mantém, e pode descer a qualquer momento.

■ Agente, assistente e automação

Assistentes agênticos, tratados em outra apostila desta coleção, são ferramentas usadas por uma pessoa que acompanha o trabalho. Agentes, aqui, são sistemas que executam uma função com menos acompanhamento, frequentemente disparados por eventos, como a chegada de um e-mail, ou por horários. Automações seguem passos fixos. Na prática, as melhores soluções combinam os três.

O que delegar primeiro

Comece onde o ganho é claro e o erro é barato.

■ A régua de duas perguntas

Para cada tarefa candidata, responda a duas perguntas, com notas de 1 a 5:

1. **Quanto tempo esta tarefa consome por semana?** (valor da delegação)
2. **Quanto dano um erro nesta tarefa pode causar?** (custo do erro)

Tarefas com muito tempo e pouco dano são as primeiras candidatas. Tarefas com muito tempo e muito dano podem ser delegadas no nível 1 ou 2, com aprovação. Tarefas com pouco tempo raramente justificam um agente.



Figura 3.2 — Matriz de delegação

■ Dez tarefas que costumam funcionar bem

A lista abaixo reúne tarefas que, em muitas empresas, combinam volume, padrão e baixo custo de erro, desde que configuradas com os cuidados desta apostila.

Tabela 3.1 — Dez tarefas candidatas à delegação a agentes

Tarefa	Nível inicial sugerido	Por que funciona
Classificar e priorizar mensagens recebidas	1 ou 2	Alto volume, critérios claros, erro fácil de corrigir
Resumir reuniões e listar tarefas combinadas	2	Economiza tempo, resultado revisado por quem participou
Pesquisar empresas e pessoas antes de reuniões	2	Trabalho de leitura, com fontes verificáveis
Preparar a agenda e os materiais da semana	2	Consolida informações espalhadas
Atualizar o CRM a partir de e-mails e reuniões	2	Tarefa odiada, com grande impacto na qualidade dos dados
Rascunhar respostas a dúvidas frequentes	2, depois 3	Base de respostas aprovadas permite verificar
Montar relatórios recorrentes	2, depois 4	Fontes e formato estáveis
Monitorar concorrentes e notícias do setor	3	Leitura de fontes públicas, sem ação externa
Conferir documentos contra regras	2	Aponta divergências para uma pessoa decidir
Organizar arquivos e nomear documentos	3	Reversível, com histórico



Figura 3.1 – Agentes funcionam melhor como ferramentas especializadas: uma função, bem feita

Tarefas que pedem cautela

Algumas tarefas parecem boas candidatas, mas exigem maturidade antes de qualquer autonomia: enviar mensagens a clientes sem revisão, negociar preços, efetuar pagamentos, alterar registros contábeis ou fiscais, publicar conteúdo em nome da empresa, tomar decisões sobre pessoas, como seleção e avaliação de colaboradores, e qualquer tarefa que envolva dados pessoais sensíveis. Nessas, o agente pode no máximo sugerir ou preparar, e o caso deve ser avaliado inclusive sob o ponto de vista legal.

! ATENÇÃO

Decisões sobre pessoas

O uso de IA em recrutamento, avaliação de desempenho, concessão de crédito e outras decisões que afetam pessoas exige atenção especial à LGPD, que garante ao titular o direito de solicitar revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses. Nesses casos, mantenha decisão humana, critérios documentados e possibilidade de revisão.

O briefing de um agente

Um agente é tão bom quanto o briefing que recebeu.

■ As partes do briefing

O briefing de um agente cumpre o papel da descrição de cargo e do manual de integração de um colaborador. Ele deve ser escrito antes da configuração técnica, revisado pelo dono do processo e mantido junto às instruções do agente.

Tabela 4.1 — Estrutura do briefing de um agente

Seção	Pergunta que responde
Função	Qual é a única função deste agente?
Objetivo e critério de sucesso	Como saber que o trabalho foi bem feito?
Gatilho	O que dispara a execução: evento, horário ou pedido?
Entradas	Quais dados recebe e de onde?
Ferramentas e permissões	O que pode consultar, o que pode alterar, o que nunca pode fazer?
Passos e regras	Há uma ordem? Quais regras de negócio se aplicam?
Formato da saída	Como entregar o resultado e onde?
Verificações	O que conferir antes de entregar?
Escalonamento	Quando parar e chamar uma pessoa? Quem?
Limites de recursos	Tempo máximo, número de passos, orçamento de uso
Registros	O que deve ficar anotado em cada execução?
Responsável	Quem responde pelo agente e revisa os indicadores?

■ Um exemplo completo

AGENTE: Pesquisador de contas
Função: preparar o dossiê de uma empresa antes da primeira reunião comercial.
Gatilho: reunião marcada no CRM com etapa "primeiro contato".
Entradas: nome e site da empresa; nome e cargo do contato; notas do vendedor.
Ferramentas: busca na web (somente leitura); leitura do CRM; escrita na pasta dossies/. Não pode enviar mensagens nem alterar registros do CRM.
Passos:
1. Ler o site da empresa e as notas do vendedor.
2. Pesquisar notícias dos últimos 12 meses em fontes confiáveis.
3. Identificar segmento, porte aproximado, produtos e possíveis necessidades relacionadas às nossas ofertas (ver empresa/ofertas.md).
4. Sugerir três perguntas para a reunião.
Saída: dossiê de uma página em dossies/<empresa>.md, com fontes de cada fato.
Verificações: todo fato tem link de origem; nada sobre a vida pessoal do contato;
informações incertas marcadas como "a confirmar".
Escalonamento: se a empresa for concorrente, cliente atual ou órgão público, não gerar dossiê; avisar o vendedor.
Limites: até 15 páginas consultadas e 10 minutos por execução.
Responsável: coordenação comercial. Revisão semanal dos indicadores.

Repare que o exemplo define o que **não** pesquisar: informações pessoais do contato. Além de uma questão de ética, essa limitação reduz riscos relacionados à LGPD, já que dados pessoais coletados sem necessidade para a finalidade comercial podem configurar tratamento excessivo.

■ As instruções de sistema, comentadas

O briefing é escrito para pessoas. A partir dele, monta-se a **instrução de sistema** do agente, o texto que o modelo recebe em toda execução. Um bom modelo de instrução tem blocos previsíveis:

1. **Papel e contexto:** “Você é o pesquisador de contas da empresa X, que vende Y para Z.” Uma ou duas frases bastam; o restante do contexto vem dos arquivos da empresa.
2. **Objetivo e critério de sucesso:** o que o resultado deve permitir e como saber que está bom.

3. **Processo:** passos na ordem em que importam, sem detalhar o que o modelo já sabe fazer.
4. **Regras e limites:** o que nunca fazer e o que exige escalonamento, com o motivo de cada regra.
5. **Uso das ferramentas:** quando usar cada uma e o que fazer quando uma falhar.
6. **Formato da saída:** estrutura exata, com um exemplo curto.
7. **Verificação final:** a lista de conferência que o agente aplica antes de entregar.

Evite repetir regras em vários blocos, usar linguagem enfática para tudo e incluir instruções que não se aplicam a esta função. Instruções enxutas são mais fáceis de manter, mais baratas por execução e produzem comportamento mais estável.

■ Ferramentas bem desenhadas

As ferramentas que o agente usa também precisam de projeto. A Anthropic publicou orientações sobre como escrever ferramentas eficazes para agentes, e a experiência prática confirma alguns princípios:

- **nomes e descrições claros:** o modelo escolhe a ferramenta pela descrição, então ela deve dizer o que faz, quando usar e o que retorna;
- **poucas ferramentas, bem delimitadas:** dez ferramentas parecidas confundem mais do que três bem definidas;
- **entradas validadas:** a ferramenta rejeita parâmetros inválidos com uma mensagem que ajuda o agente a corrigir;
- **saídas enxutas:** retornar apenas os campos necessários reduz custo e distração;
- **erros informativos:** “cliente não encontrado; verifique o CNPJ” ajuda mais do que “erro 404”;
- **ações perigosas separadas:** consultar e alterar ficam em ferramentas distintas, para que as permissões possam ser diferentes.

■ Uma função por agente

Uma regra que aparece em praticamente todas as organizações de agentes bem-sucedidas é a especialização: cada agente faz uma coisa, bem feita. Agentes generalistas, com dezenas de ferramentas e instruções para muitas situações, ficam mais imprevisíveis, mais caros e mais difíceis de avaliar. Quando uma função cresce, divida-a em dois agentes encadeados.



DICA PRÁTICA

Nomes ajudam, mas não substituem regras

Dar nomes aos agentes, como “Pesquisador” ou “Redator de propostas”, ajuda a equipe a entender quem faz o quê e a falar sobre eles como parte do fluxo de trabalho. Mas o nome não altera o comportamento: o que define o agente é o briefing.

Equipes de agentes: padrões de orquestração

Mais agentes não significa mais resultado. Significa mais coordenação.

■ Quando um agente não basta

Algumas tarefas ficam melhores divididas em etapas especializadas. Pesquisar uma empresa e escrever uma mensagem personalizada, por exemplo, são trabalhos diferentes: o primeiro exige busca e verificação de fontes; o segundo, domínio da voz da empresa e do contexto comercial. Dividir em dois agentes encadeados torna cada etapa mais simples de configurar e de avaliar.

O texto da Anthropic sobre agentes eficazes descreve padrões recorrentes de composição. Eles servem como vocabulário para desenhar equipes de agentes:

Tabela 5.1 – Padrões de orquestração de agentes

Padrão	Como funciona	Quando usar
Cadeia	A saída de uma etapa é a entrada da seguinte, com verificações entre elas	Tarefas com etapas claras e sequenciais
Roteamento	Um classificador decide qual agente ou fluxo trata cada caso	Entradas variadas que pedem tratamentos diferentes
Paralelização	Várias etapas independentes rodam ao mesmo tempo; os resultados são combinados	Análises por vários ângulos ou grandes volumes
Coordenador e executores	Um agente divide o trabalho e distribui subtarefas a outros	Tarefas complexas cujas partes não são conhecidas de antemão
Avaliador e otimizador	Um agente produz, outro avalia com critérios e pede ajustes	Quando há critérios claros de qualidade e ganho com iteração

● FLUXOGRAMA

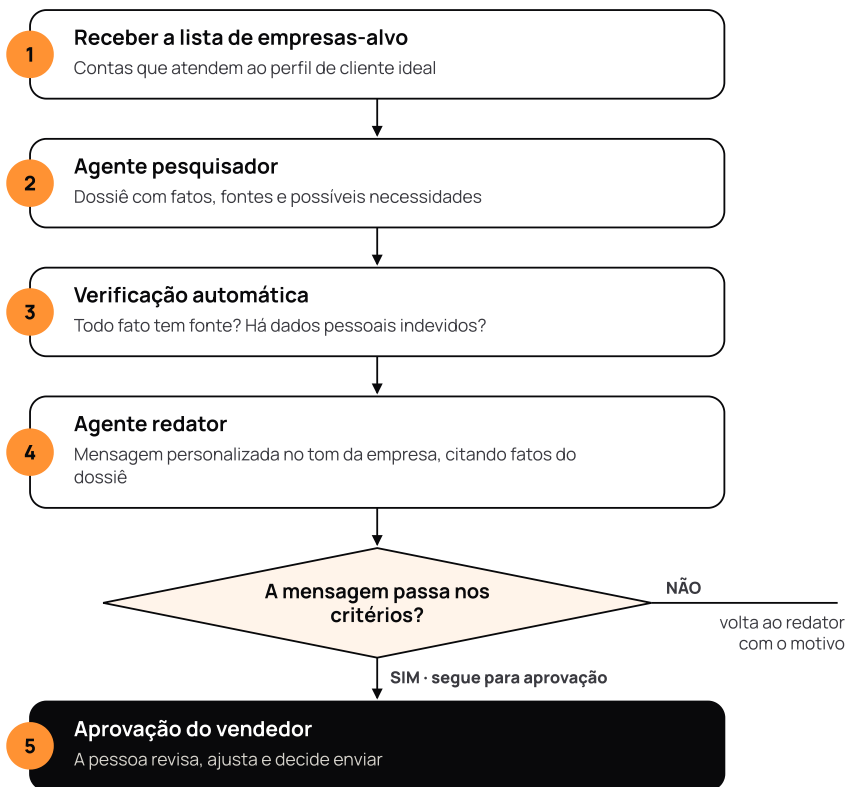


Figura 5.2 – Cadeia de dois agentes para prospecção, com verificação e aprovação

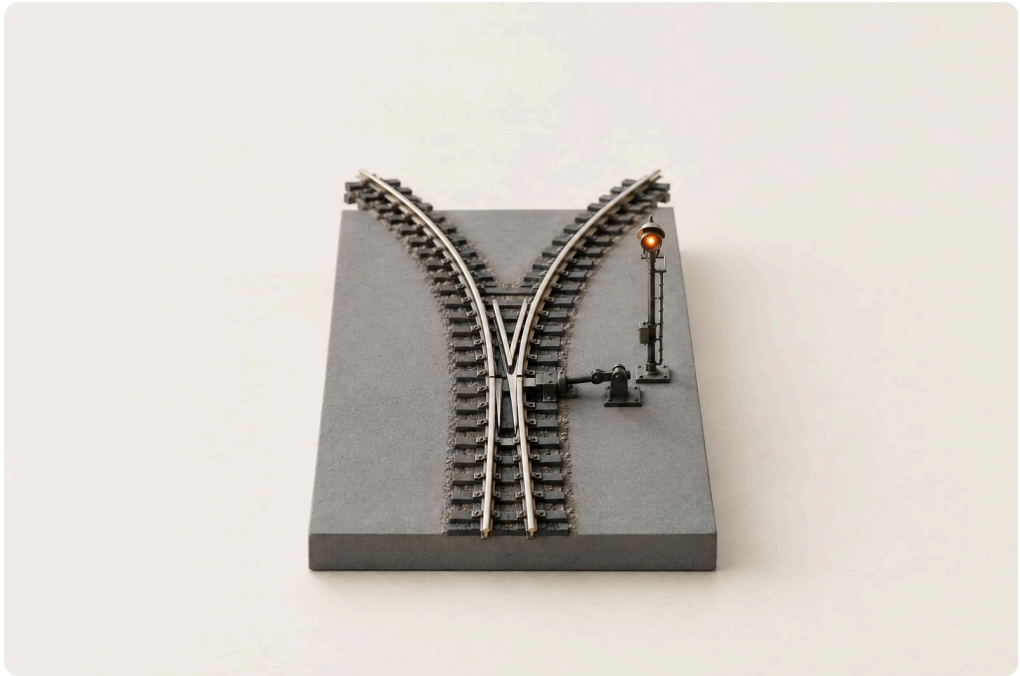


Figura 5.1 — Orquestrar é decidir, para cada caso, qual caminho seguir e quem cuida dele

Organizações de agentes por departamento

Bibliotecas abertas sobre IA aplicada a negócios têm publicado mapas de “organizações de agentes”: dezenas de agentes distribuídos por equipes de vendas, marketing, operações e finanças, cada um com uma função. Esses mapas são valiosos como inventário de possibilidades, porque mostram a variedade de tarefas que podem ser apoiadas.

Na prática, porém, poucas empresas precisam de dezenas de agentes, e nenhuma deveria começar assim. Cada agente é um sistema a ser configurado, testado, monitorado e mantido. A forma sustentável de chegar a uma equipe de agentes é incremental: um agente por vez, cada um com briefing, indicadores e responsável, e novos agentes acrescentados quando os anteriores estiverem estáveis.



Figura 5.3 — Estrutura de uma equipe de agentes madura

■ Erros comuns em equipes de agentes

- **Coordenador sobrecarregado:** um agente central com instruções para tudo, que acaba sendo um generalista frágil.
- **Responsabilidade difusa:** quando um resultado sai errado, ninguém sabe qual agente errou, porque não há registros por etapa.
- **Custo multiplicado:** cada agente consome chamadas ao modelo; cadeias longas e repetições podem multiplicar o custo sem melhorar o resultado.
- **Erros que se propagam:** um fato errado na primeira etapa contamina todas as seguintes, por isso as verificações entre etapas são essenciais.

Ferramentas e plataformas

Escolha a ferramenta pelo nível de controle de que você precisa.

■ Categorias de ferramentas

O mercado de ferramentas para agentes muda rapidamente. Em vez de uma lista de produtos que envelhece em meses, vale conhecer as categorias e o que esperar de cada uma.

Tabela 6.1 — Categorias de ferramentas para construir agentes

Categoria	Exemplos de tipo	Pontos fortes	Pontos de atenção
Plataformas de agentes sem código	Construtores visuais de agentes e de automações com IA	Rapidez, interface amigável	Dependência da plataforma, limites de controle e de custos
Plataformas de automação com IA	Ferramentas de fluxo com etapas de modelo de linguagem	Integração com muitos sistemas, fluxos previsíveis	Agentes complexos ficam difíceis de manter
Kits de desenvolvimento de agentes	Bibliotecas oficiais dos fornecedores de modelos e estruturas de código aberto	Controle total, testes, versionamento	Exigem equipe técnica
Assistentes agênticos programáveis	Ferramentas de linha de comando que executam tarefas a partir de instruções e habilidades	Reaproveitam pastas, instruções e habilidades	Mais indicados para tarefas supervisionadas
Agentes com computador próprio	Agentes que operam navegador e aplicativos em máquinas na nuvem	Funcionam onde não há API	Mais lentos, mais frágeis, exigem cuidado com credenciais

Entre os kits de desenvolvimento, estão as bibliotecas oficiais de agentes de fornecedores como OpenAI e Anthropic e estruturas abertas de orquestração, como o LangGraph. Para conectar agentes às ferramentas da empresa, o **Model Context Protocol** se consolidou como padrão aberto adotado por diversos fornecedores, o que reduz a dependência de integrações proprietárias.

■ Critérios de escolha



CHECKLIST

Perguntas para avaliar uma ferramenta de agentes

- ☐ É possível limitar as permissões de cada agente a ações específicas?
- ☐ Há registro detalhado de cada execução, com entradas, passos e saídas?
- ☐ É possível exigir aprovação humana antes de ações selecionadas?
- ☐ Há controle de custos, com limites por agente e alertas?
- ☐ É possível testar o agente com casos de teste antes de publicar mudanças?
- ☐ As instruções e configurações podem ser exportadas e versionadas?
- ☐ Os dados ficam em qual país, por quanto tempo e sob qual contrato?
- ☐ A ferramenta permite trocar o modelo de linguagem usado?

■ Estimando o custo de um agente

O custo de um agente baseado em modelos de linguagem depende principalmente da quantidade de texto processado, medida em tokens, e do preço do modelo. Cada execução soma as instruções de sistema, os documentos e resultados de ferramentas lidos, os passos de raciocínio e o texto produzido. Agentes que fazem muitas consultas ou leem documentos longos consomem mais.

E EXEMPLO PRÁTICO**Estimativa de custo mensal de um agente**

Um agente de pesquisa de contas executa 400 vezes por mês. Cada execução processa, em média, cerca de 30 mil tokens de entrada, somando instruções, páginas lidas e resultados de ferramentas, e produz 2 mil tokens de saída. O custo mensal é aproximadamente: $400 \times (30.000 \times \text{preço por token de entrada} + 2.000 \times \text{preço por token de saída})$.

Com os preços vigentes do modelo escolhido, a conta mostra rapidamente se o agente compensa. Se o custo por execução for pequeno diante do tempo economizado, o caso se sustenta. Se não, há alavancas: modelo mais econômico, menos páginas lidas por execução, saídas mais curtas ou reaproveitamento de resultados. (Raciocínio ilustrativo; consulte a tabela de preços atual do fornecedor.)

Dois fatores costumam ser esquecidos nas estimativas: as **repetições**, quando o agente refaz passos após erros de ferramentas, e o **custo humano de revisão**, que deve entrar na conta do retorno.

Supervisão, avaliação e segurança

Um agente sem medição é uma aposta. Com medição, é um processo.

■ Registros de execução

Cada execução de um agente deve deixar um rastro: o que recebeu, quais ferramentas usou, quais resultados obteve, o que decidiu e o que entregou. Esses registros, chamados de rastros ou *traces*, são indispensáveis para investigar erros, medir custos e melhorar instruções. Sem eles, a pergunta “por que o agente fez isso?” não tem resposta.

■ Avaliação antes e depois de publicar

A avaliação de agentes combina duas práticas:

- **Avaliação offline:** um conjunto de casos de teste, com entradas reais e resultados esperados ou critérios, executado sempre que as instruções, as ferramentas ou o modelo mudarem.
- **Avaliação online:** acompanhamento do trabalho real, por amostragem de revisões humanas, indicadores de aprovação e alertas automáticos.

Uma técnica frequente é usar um modelo de linguagem como avaliador, comparando a saída do agente com critérios escritos. Pesquisas, como o estudo de Lianmin Zheng e colegas sobre modelos como juízes, mostram que essa avaliação pode ter boa concordância com avaliadores humanos em muitas tarefas, mas também apresenta vieses, como preferência por respostas mais longas. Por isso, o avaliador automático deve ser calibrado periodicamente com avaliações humanas.

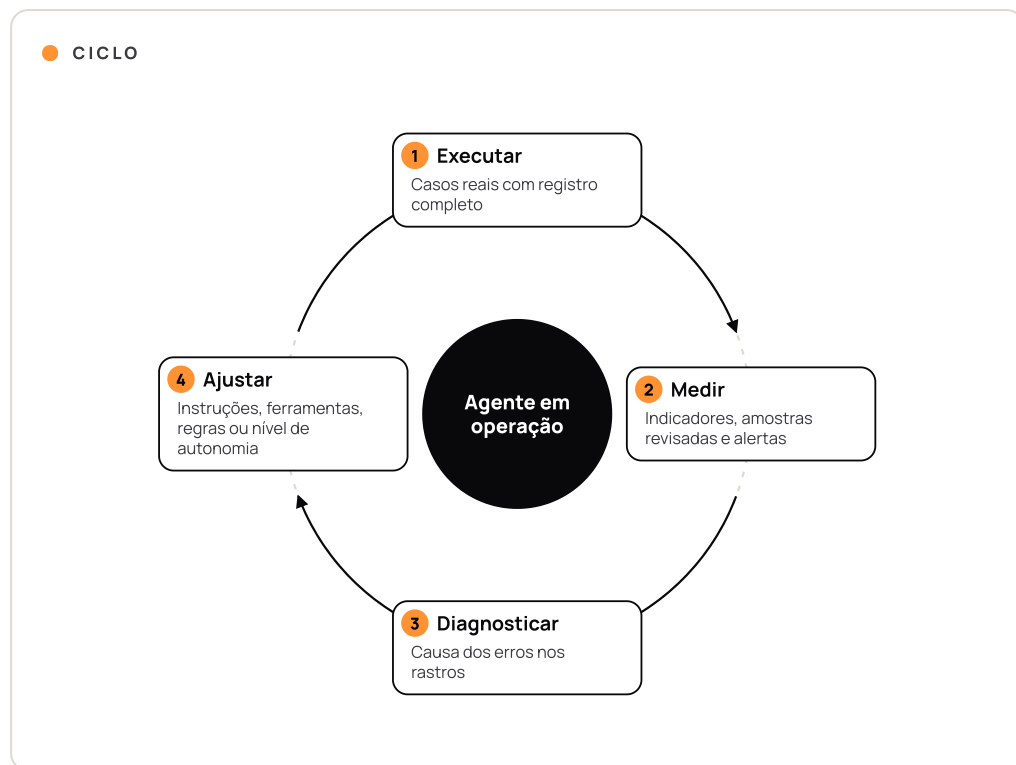


Figura 7.1 – Ciclo de avaliação e melhoria de um agente

■ Riscos de segurança

Tabela 7.1 — Principais riscos de segurança de agentes e controles recomendados

Risco	Como acontece	Controles
Injeção de instruções	Conteúdo lido pelo agente traz ordens escondidas	Separar leitura de ação, aprovação humana, tratar conteúdo externo como dado
Autonomia excessiva	Permissões e funções além do necessário	Privilegio mínimo, listas de ações permitidas, níveis de autonomia
Exposição de dados	Agente revela ou envia dados indevidos	Filtros de permissão, minimização de dados, revisão de saídas externas
Consumo sem limite	Repetições e cadeias longas geram custo alto	Limites de passos, tempo e orçamento, alertas
Erros que se propagam	Fato errado contamina etapas seguintes	Verificações entre etapas, fontes obrigatórias
Credenciais expostas	Senhas em instruções ou ao alcance do agente	Cofres de segredos, credenciais por agente com escopo mínimo

Esses riscos correspondem a itens da lista do OWASP para aplicações com modelos de linguagem, como injeção de instruções, divulgação de informações sensíveis, autonomia excessiva e consumo sem limites.

■ Regras aprendidas da forma difícil

! ATENÇÃO

Regras práticas de segurança para agentes

1. Nenhum agente recebe, no primeiro dia, permissão para enviar mensagens externas, pagar ou apagar.
2. Agentes que leem conteúdo da internet ou de terceiros não executam ações sensíveis na mesma sessão sem aprovação.
3. Cada agente tem credenciais próprias, com o menor escopo possível, revogáveis a qualquer momento.
4. Todo agente tem limite de passos, de tempo e de custo por execução.
5. Existe um botão de parada: qualquer responsável pode suspender o agente imediatamente.
6. Mudanças de instrução ou de modelo passam pelos casos de teste antes de entrar em operação.

■ Quando algo dá errado: resposta a incidentes

Mesmo com todos os cuidados, incidentes acontecem: um agente envia um rascunho ao destinatário errado, consome orçamento em repetições, altera um registro indevidamente. Ter um roteiro definido antes do problema reduz o dano.

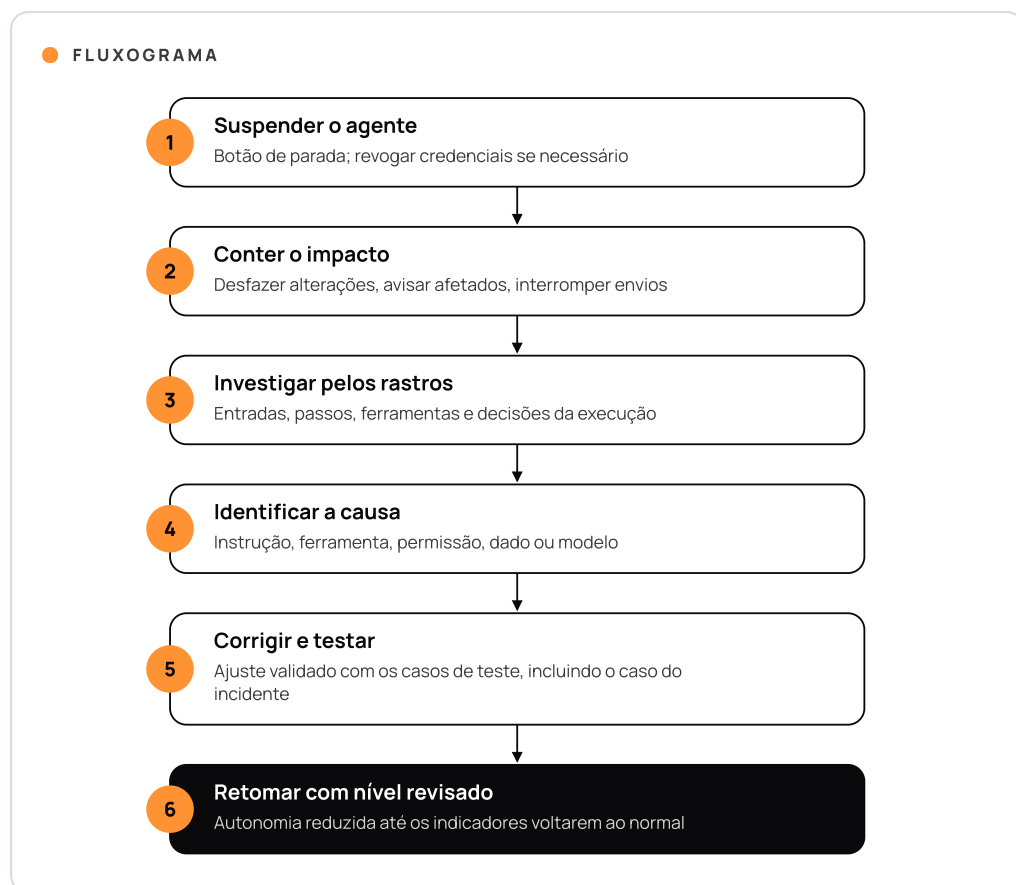


Figura 7.2 – Roteiro de resposta a incidentes com agentes

Se o incidente envolver dados pessoais, avalie também as obrigações previstas na LGPD, que incluem a comunicação à Autoridade Nacional de Proteção de Dados e aos titulares em casos de incidente de segurança que possa acarretar risco ou dano relevante. O encarregado de dados da empresa deve participar da análise.

Implantação e indicadores

Oito semanas para o primeiro agente; a rotina de revisão para os seguintes.

A implantação do primeiro agente funciona como um projeto-piloto da própria capacidade da empresa de trabalhar com agentes. Além de entregar a tarefa automatizada, ela testa o processo de escolha, o briefing, a configuração de permissões, os casos de teste, a revisão humana e a rotina de indicadores. Por isso, vale escolher uma tarefa de valor claro e baixo risco e documentar cada decisão: os agentes seguintes vão reaproveitar esse caminho, com muito menos esforço.



Figura 8.1 — Implantação do primeiro agente em oito semanas

Tabela 8.1 — Indicadores para agentes de IA

Indicador	Definição	Uso
Taxa de sucesso	Execuções concluídas corretamente + execuções	Qualidade geral
Taxa de aprovação sem alteração	Saídas aprovadas sem mudança + saídas revisadas	Decidir sobre subir de nível
Taxa de escalonamento	Casos enviados a pessoas + casos recebidos	Verificar se o agente reconhece seus limites
Severidade dos erros	Erros classificados como leves, moderados e graves	Decidir sobre descer de nível
Tempo economizado	Tempo da linha de base menos tempo atual, incluindo revisão	Retorno do agente
Custo por execução	Consumo de modelo e ferramentas + execuções	Sustentabilidade
Tempo de ciclo	Tempo entre o gatilho e a entrega	Experiência do cliente interno ou externo

■ O ritual semanal de revisão

Agentes em operação precisam de uma reunião curta e regular, de 20 a 30 minutos, com o dono do processo e o responsável técnico. Uma pauta fixa mantém o foco:

- 1. Indicadores da semana**, comparados à semana anterior e à linha de base.
- 2. Erros e escalonamentos**: três ou quatro casos examinados pelos rastros.
- 3. Mudanças propostas**: instruções, ferramentas, regras ou nível de autonomia, sempre com casos de teste.
- 4. Custos**: consumo da semana e eventuais anomalias.
- 5. Decisões**: o que muda, quem faz e quando será verificado.

Essa rotina transforma o agente em um processo gerenciado. Sem ela, a qualidade tende a se degradar silenciosamente, até que um erro visível provoque a desconfiança da equipe.

**DICA PRÁTICA****Regra para mudar de nível**

Defina por escrito, antes do piloto, as condições para mudar o nível de autonomia. Por exemplo: subir do nível 2 para o 3 depois de quatro semanas com aprovação sem alteração acima de 95% e nenhum erro grave; voltar ao nível 2 diante de qualquer erro grave.

Estudo de caso ilustrativo e exercícios



ESTUDO DE CASO ILUSTRATIVO

AgroPeças Cerrado: três agentes e uma aprovação

A AgroPeças Cerrado distribui peças para máquinas agrícolas e recebe, por e-mail e mensagens, cerca de 70 pedidos de cotação por dia. Três atendentes liam cada pedido, identificavam as peças, consultavam estoque e preço no sistema de gestão e redigiam a cotação. Nos períodos de safra, o tempo de resposta passava de dois dias, e clientes compravam da concorrência.

Desenho. A equipe adotou uma cadeia de três agentes. O primeiro lê o pedido e extrai as peças, quantidades e modelo da máquina, consultando um dicionário de nomes populares de peças. O segundo consulta estoque e preço no sistema de gestão, com acesso somente leitura. O terceiro redige a cotação no modelo da empresa. Uma atendente aprova cada cotação antes do envio.

Regras e escalonamento. Pedidos com peças não identificadas, quantidades fora do padrão ou clientes com pendência financeira vão direto para uma pessoa, com o motivo indicado. O agente redator não pode oferecer descontos; ele apenas aplica a tabela vigente.

Verificações. Entre o segundo e o terceiro agente, uma verificação automática compara cada item com o cadastro e impede cotações com preço zerado ou estoque negativo.

Piloto. Nas duas primeiras semanas, cerca de um quarto das cotações precisou de ajuste, a maioria por peças com nomes regionais que não estavam no dicionário. Cada ajuste alimentou o dicionário. A empresa passou a acompanhar a taxa de aprovação sem alteração, o tempo entre o pedido e o envio e o percentual de pedidos escalados.

Decisão. Com os indicadores estabilizados, a diretoria decidiu manter a aprovação humana para todas as cotações, priorizando a relação com os clientes, e usar o tempo liberado das atendentes para retornar aos clientes que não responderam às cotações.

Caso ilustrativo: empresa, pessoas e números são fictícios, criados para fins didáticos. Não representam clientes nem resultados da D.B.

**EXERCÍCIO DE APLICAÇÃO****Exercício 1: matriz de delegação**

1. Liste dez tarefas recorrentes da sua área.
2. Dê a cada uma nota de 1 a 5 para tempo consumido e para custo de um erro.
3. Posicione as tarefas na matriz do capítulo 3 e escolha a primeira candidata.

**EXERCÍCIO DE APLICAÇÃO****Exercício 2: briefing completo**

1. Escreva o briefing do agente para a tarefa escolhida, com todas as seções do capítulo 4.
2. Defina o nível de autonomia inicial e as condições para mudar de nível.
3. Liste dez casos de teste, incluindo três exceções.

**EXERCÍCIO DE APLICAÇÃO****Exercício 3: mapa de riscos**

1. Para o agente desenhado, preencha a tabela de riscos do capítulo 7 com os controles que você usaria.
2. Identifique quais ações dependem de aprovação humana.
3. Descreva como alguém da equipe suspenderia o agente em caso de problema.

Como a D.B. pode atuar e conclusão

■ Conclusão

Agentes de IA podem assumir uma parte relevante do trabalho repetitivo das empresas, mas só entregam valor quando são tratados como membros de uma equipe que precisam de função clara, instruções, ferramentas adequadas, supervisão e avaliação. A estratégia que funciona é gradual: escolher tarefas com muito tempo e erro barato, começar com níveis baixos de autonomia, medir contra uma linha de base e subir de nível apenas com dados.

Equipes de agentes, com especialistas por departamento, são um destino possível, não um ponto de partida. Cada agente acrescentado deve ter dono, indicadores e espaço na rotina de revisão. Assim, a empresa constrói uma capacidade real, e não uma coleção de demonstrações.

D COMO A D.B. PODE AJUDAR

Como a D.B. implanta agentes de IA

A D.B. Soluções Digitais desenha e implanta agentes com função clara e supervisão humana:

- seleção das tarefas com a matriz de delegação e linha de base;
- briefing, instruções, ferramentas com permissões mínimas e casos de teste;
- orquestração com verificações entre etapas e escalonamento para pessoas;
- registros de execução, controle de custos e indicadores;
- revisão periódica para ajustar instruções e níveis de autonomia.

Glossário

Agente de IA

Sistema que persegue um objetivo decidindo passos e usando ferramentas, com base em um modelo de linguagem.

Avaliação offline

Teste do agente com casos conhecidos antes de publicar mudanças.

Avaliação online

Acompanhamento da qualidade do agente durante a operação real.

Briefing de agente

Documento que define função, entradas, ferramentas, regras, saídas, verificações, escalonamento e limites de um agente.

Escalonamento

Encaminhamento de um caso do agente para uma pessoa, segundo regras definidas.

Modelo como juiz

Uso de um modelo de linguagem para avaliar a saída de outro, com critérios escritos.

Nível de autonomia

Grau de independência concedido a um agente, da sugestão à execução autônoma.

Orquestração

Coordenação de vários agentes e etapas para executar uma tarefa maior.

Rastro (trace)

Registro detalhado das entradas, passos, chamadas de ferramentas e saídas de uma execução.

Roteamento

Padrão em que um classificador decide qual fluxo ou agente trata cada caso.

REFERÊNCIAS

Referências

As referências abaixo fundamentam conceitos, normas e boas práticas citados no texto. A organização temática da coleção tomou como ponto de partida a biblioteca aberta Altari AI Guide (altari.ai/guide); nenhum texto, imagem ou diagrama dela foi reproduzido. Ferramentas e produtos de terceiros mudam com frequência: confira a documentação oficial antes de implementar. Links verificados em outubro de 2026.

1. ALTARI. *The AI Teammate Playbook, The 16-Agent AI Workforce, The 37-Agent Marketing Org, The 42-Agent Sales Org e The 2-Agent Outbound System*. Altari AI Guide, 2026. Disponível em: <https://altari.ai/guide>. Referência temática; acesso em: 21 out. 2026.
2. ANTHROPIC. *Building effective agents*. 19 dez. 2024. Disponível em: <https://www.anthropic.com/engineering/building-effective-agents>.
3. ANTHROPIC. *Writing effective tools for agents – with agents*. 2025. Disponível em: <https://www.anthropic.com/engineering/writing-tools-for-agents>.
4. ANTHROPIC. *Agent SDK overview*. Documentação oficial. Disponível em: <https://docs.claude.com/en/docs/agent-sdk/overview>.
5. BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD), art. 20. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm.
6. LANGCHAIN. *LangGraph*. Documentação. Disponível em: <https://langchain-ai.github.io/langgraph/>.
7. MODEL CONTEXT PROTOCOL. *Introduction*. Disponível em: <https://modelcontextprotocol.io/>.
8. NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1, 2023. Disponível em: <https://doi.org/10.6028/NIST.AI.100-1>.
9. OPENAI. *A practical guide to building agents*. 2025. Disponível em: <https://cdn.openai.com/business-guides-and-resources/a-practical-guide-to-building-agents.pdf>.
10. OPENAI. *OpenAI Agents SDK*. Documentação. Disponível em: <https://openai.github.io/openai-agents-python/>.
11. OWASP FOUNDATION. *OWASP Top 10 for LLM Applications 2025*. Disponível em: <https://genai.owasp.org/llm-top-10/>.
12. PARASURAMAN, Raja; SHERIDAN, Thomas B.; WICKENS, Christopher D. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A*, v. 30, n. 3, p. 286-297, 2000.
13. RUSSELL, Stuart; NORVIG, Peter. *Inteligência artificial: uma abordagem moderna*. 4. ed. Rio de Janeiro: GEN LTC, 2022.
14. WOOLDRIDGE, Michael. *An introduction to multiagent systems*. 2. ed. Chichester: Wiley, 2009.
15. YAO, Shunyu et al. ReAct: synergizing reasoning and acting in language models. In: *ICLR*, 2023. Disponível em: <https://arxiv.org/abs/2210.03629>.
16. ZHENG, Lianmin et al. Judging LLM-as-a-judge with MT-Bench and Chatbot Arena. In: *NeurIPS*, 2023. Disponível em: <https://arxiv.org/abs/2306.05685>.

— PRÓXIMO PASSO

Da leitura à *implementação*.

A D.B. desenha e implanta agentes com função clara, ferramentas controladas e supervisão humana, começando pela tarefa de maior retorno e menor risco.

Diagnóstico

Mapeamos processos, dados e ferramentas para encontrar onde a IA e a automação fazem sentido.

Automações e integrações

Conectamos sistemas e criamos fluxos com regras claras, tratamento de falhas e registro.

IA aplicada

Assistentes, agentes e bases de conhecimento com limites definidos e revisão humana.

Sistemas e sites

Desenvolvimento sob medida, com foco em uso real, desempenho e evolução.

Conversa inicial sem compromisso: você explica o contexto, nós ajudamos a enxergar o caminho.

D.B. Soluções Digitais

Além do código.

WhatsApp (98) 99107-0234

Conversar com a D.B. ↗