

ARSENAL DO DETETIVE

ESPECIALIDADE

OSINT com agentes de IA

Configure, supervisione e avalie pesquisas com fontes verificáveis.

Explicações / Exemplos / Modelos para adaptar

Conhecimento que acompanha o seu trabalho.

Edição 5 · Setembro de 2026

Leitura e prática para o investigador profissional.

Licença individual. Exemplos de treinamento identificados.

14. Caso documentado: instruções escondidas em páginas da web**32****Fontes para continuar estudando****35**

Aprenda a organizar uma pesquisa, usar uma automação pequena e revisar o resultado antes de entregá-lo ao cliente.

1. A primeira ferramenta é uma pergunta bem feita

Um cliente está avaliando um fornecedor de software. Na apresentação comercial, encontrou uma referência à NASA. Ele quer saber se isso significa que o fornecedor é um parceiro oficial da agência americana. A frase impressiona, mas não explica a relação. O fornecedor pode utilizar um programa público desenvolvido pela instituição; pode ter participado de um contrato; pode apenas mencionar uma tecnologia conhecida. São situações diferentes.

Antes de abrir uma ferramenta de inteligência artificial, o investigador precisa entender qual dessas afirmações foi feita e qual decisão o cliente pretende tomar. Se o objetivo é decidir se vale a pena pedir uma proposta detalhada, uma verificação limitada de fontes públicas pode ajudar. Se o cliente pretende assinar um contrato de grande valor, a mesma pesquisa talvez sirva apenas para identificar documentos e confirmações adicionais que serão necessários.

Neste livro, vamos trabalhar com uma situação comercial inventada e um projeto público real: o Open MCT, divulgado pela NASA. A empresa do exemplo será chamada de Estúdio Horizonte. Ela é fictícia; não estamos avaliando uma organização existente. O projeto público serve para ensinar a leitura de fontes. Nenhuma instituição citada apoia este material, certifica o nosso método ou recomenda os produtos da loja.

A pergunta do exercício será: “As fontes examinadas identificam o Open MCT como um projeto relacionado à NASA e demonstram que o Estúdio Horizonte é parceiro oficial da agência?”. A primeira parte pode encontrar apoio em páginas institucionais. A segunda exige uma ligação específica com o fornecedor. Um agente de IA precisa respeitar essa diferença. Você também.

O que significa fazer OSINT

OSINT é a sigla inglesa usada para o trabalho de produzir informação útil a partir de fontes abertas. Uma fonte aberta pode ser uma página institucional, um documento publicado, um registro acessível ao público ou um comunicado oficial. O trabalho envolve escolher o que consultar, verificar a informação, compreender seu contexto e explicar o resultado.

Encontrar uma página não encerra a pesquisa. Você precisa saber quem publicou, a que assunto o texto se refere, quando a informação foi observada e se ela realmente responde à pergunta. Um comunicado sobre um projeto da NASA não identifica automaticamente todas as empresas que usam esse projeto. Um endereço na internet contendo uma palavra famosa também não prova vínculo institucional.

O valor do investigador profissional aparece nessa leitura. Ele percebe quando uma afirmação é mais ampla do que a fonte permite e consegue explicar a diferença sem acusar alguém por antecipação. Se a fonte não esclarece uma questão, registra o limite e indica uma próxima verificação proporcional.

Na prática, uma pesquisa pequena começa com três anotações: o que o cliente perguntou, qual material originou a dúvida e qual resposta será útil. No nosso exemplo, o material inicial é uma frase comercial inventada: “Nossa solução utiliza Open MCT, tecnologia da NASA; por isso, somos parceiros oficiais da agência”. A primeira oração e a conclusão precisam ser examinadas separadamente.

Você pode escrever a pergunta em um arquivo de texto e guardar a frase recebida em outro documento identificado. O arquivo de trabalho deve informar que a frase é fornecida pelo cliente, ainda não confirmada. Isso impede que, ao copiar o texto para outra etapa, uma alegação passe a parecer um fato apurado.

Faça uma primeira pesquisa sem IA

Abra o catálogo público de software da NASA, em code.nasa.gov, e procure o nome Open MCT. A página pode mudar de organização; o objetivo é localizar a referência ao projeto e conferir o destino do link. No levantamento usado neste livro, o catálogo encaminhava para o repositório [nasa/openmct](https://github.com/nasa/openmct) no GitHub.

Um repositório é uma página organizada de um projeto de software, com arquivos, informações e histórico de desenvolvimento. Para este exercício, você não precisa programar, instalar o projeto ou ler o código. Precisa reconhecer qual organização publica a página e quais informações estão efetivamente disponíveis.

Anote o endereço institucional consultado, o destino do link e a data em que você fez a consulta. Depois, abra a página do projeto. Procure a descrição e a área de versões publicadas. Uma versão publicada, chamada de release em muitos serviços, é uma edição que os responsáveis disponibilizaram com uma identificação própria. Ela não deve ser confundida com qualquer alteração feita nos arquivos.

Ao final dessa primeira leitura, já é possível escrever uma resposta curta: “O catálogo institucional consultado identifica o projeto Open MCT e aponta para o repositório [nasa/openmct](https://github.com/nasa/openmct). Esse material estabelece o contexto do projeto. Não apresenta, por si, uma relação comercial com o fornecedor fictício citado na pergunta”.

Essa resposta não exige agente, servidor ou assinatura. Ela serve de referência para avaliar a automação que construiremos. Se a ferramenta produzir uma conclusão mais forte sem encontrar uma fonte adicional adequada, o ganho de fluência terá escondido uma perda de qualidade.

Para consultar: o [catálogo de software da NASA](https://code.nasa.gov) e a [página institucional da divisão de sistemas inteligentes](https://github.com/nasa/openmct) permitem conhecer o contexto público do projeto.

2. Entenda o que um agente faz antes de pedir autonomia

Você provavelmente já usou um assistente de IA: escreveu uma pergunta, recebeu uma resposta e pediu uma alteração. Nesse formato, a pessoa escolhe a próxima ação. Se for necessário abrir uma página, buscar um documento ou confirmar uma data, ela decide como proceder ou utiliza uma função disponível no serviço.

Uma automação fixa segue uma sequência que você preparou. Por exemplo: consultar duas páginas conhecidas, extrair alguns campos e gravar o resultado em um arquivo. A sequência pode ser executada novamente, mas não precisa decidir livremente onde pesquisar. Esse tipo de trabalho costuma ser suficiente quando a tarefa se repete e as fontes são previsíveis.

Um agente combina um modelo de IA com instruções e ferramentas. Dentro das possibilidades que você disponibiliza, o modelo pode escolher uma ferramenta, interpretar a resposta e decidir se precisa de outra consulta antes de redigir o resultado. Essa escolha é a parte que distingue o agente da sequência inteiramente predeterminada.

Imagine um auxiliar recebendo uma pasta com três documentos e uma pergunta. Você pode pedir que ele leia os três em uma ordem fixa. Também pode permitir que escolha primeiro qual documento consultar. Na segunda situação, ele tem alguma autonomia sobre o caminho; continua limitado à pasta, à pergunta e às ações permitidas.

No laboratório deste livro, o agente terá uma função semelhante. Ele poderá solicitar três fichas de fontes previamente preparadas. Não poderá escolher novos sites, abrir contas, enviar mensagens, executar programas ou acessar documentos do cliente. Essa limitação torna o exercício mais simples de compreender e ajuda a descobrir se a resposta está sustentada pelo material disponível.

O que o modelo sabe não substitui a fonte

Modelos de linguagem produzem respostas a partir do contexto recebido e do que aprenderam durante seu desenvolvimento. Eles podem explicar conceitos com clareza e, ao mesmo tempo, errar uma data, completar uma lacuna ou atribuir uma afirmação a uma página que não a contém.

Por isso, “o modelo conhece a NASA” não resolve a pergunta do cliente. Precisamos saber quais fontes foram consultadas para aquela resposta e o que elas permitem afirmar. Uma referência genérica à reputação da agência não demonstra uma relação entre a agência e um fornecedor.

Quando o agente solicita uma ficha e recebe um texto, esse texto passa a integrar o material que ele usa para responder. A ferramenta não transforma o conteúdo em verdade. Se a ficha contém uma alegação, ela continua sendo uma alegação. Se a coleta falhou, o agente precisa tratar essa ausência como uma falha de acesso, sem preencher o espaço com uma lembrança.

Também não confunda uma resposta bem organizada com uma pesquisa completa. O modelo pode produzir uma tabela convincente a partir de apenas uma fonte. A tabela melhora a apresentação; não aumenta a independência das informações que a alimentaram.

Escolha a forma mais simples que resolve o trabalho

Considere três tarefas: conferir duas datas, acompanhar periodicamente uma lista pública e comparar argumentos em vários documentos autorizados. Na primeira, a leitura manual pode ser mais rápida do que montar um fluxo. Na segunda, uma automação fixa pode registrar alterações com menos variação do que um agente. Na terceira, um modelo pode ajudar a organizar diferenças, desde que os documentos possam ser enviados ao serviço e que uma pessoa confira as referências.

O agente se torna interessante quando precisa escolher entre ferramentas conhecidas ou adaptar uma pequena sequência à pergunta. Ele se torna um problema quando a autonomia foi adicionada sem necessidade, o objetivo é vago ou ninguém consegue explicar o que foi consultado.

Necessidade	Primeira abordagem razoável	O que conferir
Ler uma informação em uma página conhecida	Consulta manual	Identidade da página, contexto e data
Recolher sempre os mesmos campos	Automação fixa	Erros de acesso, campos ausentes e alterações
Comparar documentos autorizados	Assistente com material selecionado	Correspondência entre afirmação e trecho
Escolher entre poucas consultas permitidas	Agente com ferramentas limitadas	Consultas realizadas, conclusão e ponto de parada

Essa escolha afeta o preço do serviço. Uma pesquisa que exige quinze minutos de leitura não melhora comercialmente só porque foi cercada de uma instalação complexa. O cliente paga por um resultado útil e responsável. Seu custo de preparação precisa fazer sentido para a frequência do trabalho e para o risco de erro.

Para consultar: o [perfil do NIST para IA generativa](#) discute, entre outros riscos, respostas incorretas apresentadas com confiança. A referência ajuda a compreender por que avaliação humana e qualidade das fontes continuam necessárias.

3. Combine o serviço antes de automatizar a pesquisa

A ferramenta não altera a responsabilidade assumida com o cliente. Se você recebeu documentos para verificar uma questão comercial, não pode concluir que qualquer serviço de IA está autorizado a recebê-los. Antes de utilizar uma plataforma externa, precisa avaliar a finalidade, a necessidade, a confidencialidade e as condições aplicáveis ao tratamento daquela informação.

Uma maneira simples de conversar sobre o serviço é perguntar: “Qual afirmação precisa ser conferida?”, “O que você pretende decidir com a resposta?” e “Qual documento você recebeu?”.