PORTUGAL DIGITAL SUMMIT 2026 · KEYNOTE

Quatro perguntas antes de pôr um agente em produção

Página de apoio à minha keynote no Portugal Digital Summit, a 22 de outubro de 2026. Se estiveram na sala, é o que fica depois de ela esvaziar. Se não estiveram, é a conversa por escrito.

Hélder Dias Ribeiro · 22 de outubro de 2026



Quando começámos a nossa jornada de IA, não começámos por construir agentes. Começámos pelos dados. Primeiro, uma plataforma de dados em vez de três, a um custo bastante inferior ao de manter três. Depois, governança: 2.100 entidades de dados, como vendas, stock, ruturas e preço, cada uma com dono, origem conhecida e significado acordado. Só então abrimos a porta ao negócio, com o AIOps: edições de dez semanas, com negócio e tecnologia juntos, cada uma a terminar com soluções em produção.

No último ano passámos 55 soluções de IA pelo nosso programa de adoção. Destas, 47 foram adotadas, em produção ou em rollout. Oito não foram. Só uma parou por causa do modelo. As outras sete pararam numa de quatro perguntas que hoje fazemos a todos os agentes antes de os deixarmos sair do piloto. Nenhuma das oito parou por causa de dados que já estivessem na plataforma, e a maioria das 47 adotadas corre sobre ela.

Um piloto prova que funciona uma vez. A produção tem de o provar todos os dias. As quatro perguntas são a diferença entre as duas.

As quatro perguntas

Na keynote, cada pergunta tinha uma palavra: aceder, compreender, responder e medir. Aqui a última faz-se como se faria numa reunião: compensa?

1 Consegue aceder?

O agente tem identidade própria e as permissões certas nos sistemas e dados de que precisa, para ler e para escrever, com regras escritas sobre o que pode alterar?

Aceder correu bem com os contratos comerciais. O agente tem identidade própria, separa os contratos assinados dos não assinados, prepara o email de seguimento e atualiza ele próprio o ficheiro de controlo: o único sítio onde pode escrever. Correu mal com um serviço não core que gerimos com um parceiro: os dados estavam do lado deles, e um agente nosso não entra em sistemas que não controlamos. Essa parte continua por resolver.

Um agente também precisa de um cartão de acesso.

2 Consegue compreender?

Consegue ler o que lá está, e o significado foi acordado antes de começar a ler?

Foi a pergunta que nos custou mais: quatro das oito pararam aqui. Correu bem com o inventário das lojas: um agente que combina regras fixas com raciocínio de IA para explicar porque é que uma contagem não bate certo, e acerta à primeira em quatro de cada cinco casos. Funciona porque alguém decidiu à partida o que conta como discrepância. Correu mal com a manutenção: um agente que validava pedidos de manutenção com base em relatórios de vistoria. No piloto, com ficheiros limpos e pastas dedicadas, funcionou. Em produção não havia qualquer padrão para os relatórios nem para os próprios pedidos, e é difícil pedir consistência a uma máquina num sítio onde nunca fomos consistentes. A nossa plataforma de dados tinha resolvido as tabelas. Os documentos escritos para pessoas nunca tinham sido dados. Pensem no vosso disco: quantos anos de decisões estão em apresentações e atas que ninguém voltará a abrir? As organizações escrevem hoje para dois leitores, pessoas e agentes. Para cada tipo de documento que um agente vai ler, acordamos primeiro a estrutura, e o documento novo já nasce nela. Em vários casos tivemos de a aplicar retroativamente, a documentos que já existiam.

Ninguém tinha um problema de dados até um agente precisar de os ler.

3 Quem responde por ele?

Quando o agente age, há um dono com nome que responde por ele, com regras claras definidas antes de ser construído?

O mesmo agente de inventário tem dono, o departamento de Operações, e a regra está escrita: o que não acerta à primeira vai para o responsável de stock da loja. Com o dono e as regras definidos, o plano para o alargar a toda a rede foi aprovado de uma só vez. Com um assistente de voz pensado para ligar a clientes empresariais por pagamentos em falta, só chegámos às regras depois de o construir, e nunca entrou em produção. O que nos trava não são as regras. É chegar a elas tarde.

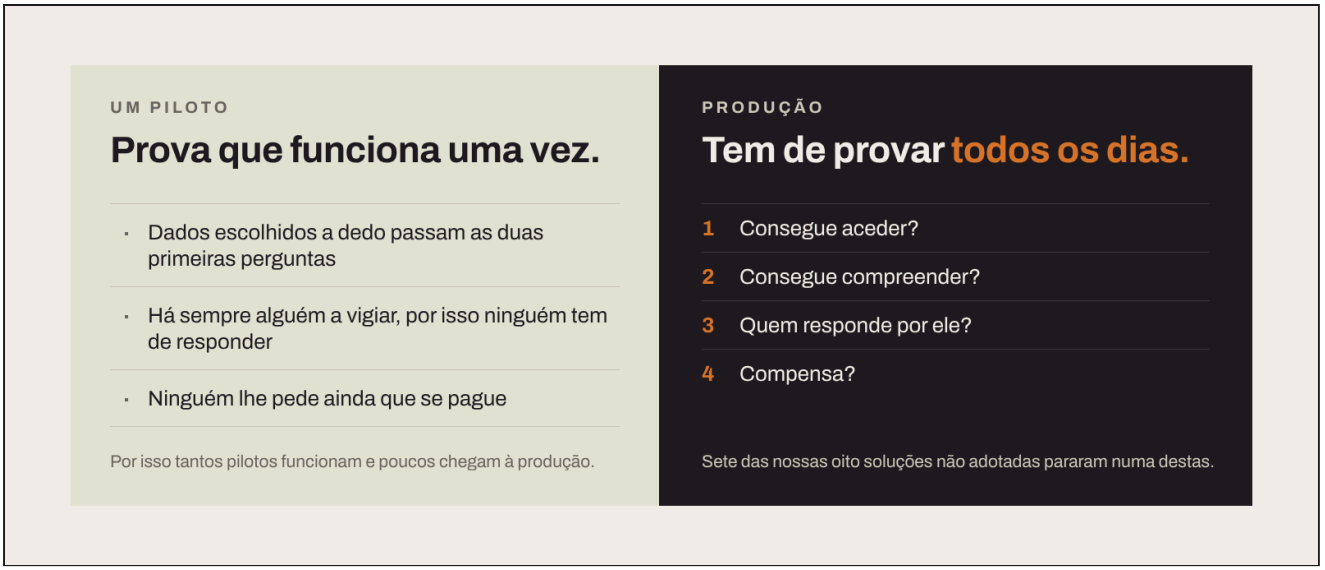
Um agente age. Uma pessoa responde por ele.

4 Compensa?

O valor que o agente cria compensa o custo de o construir e manter?

Medimos em horas e em euros, e cada solução tem de se pagar sozinha, incluindo a manutenção. Nas seis edições concluídas, as equipas estimam poder recuperar 153.000 horas por ano. Correu bem com o Storecast: um agente que escreve e lê em áudio os indicadores de cada loja, que cada diretor de loja preparava à mão para a reunião do início do dia. Poupa 15 minutos por loja por dia em cerca de 400 lojas, o que estimamos em 35.000 horas por ano. Um diretor de loja contou-me que o ouve no carro, a caminho do trabalho. Correu mal com um agente que escrevia relatórios de visita às lojas. Funcionou no piloto, e recuperaria 20 horas por ano. Não o adotámos: custaria tanto a manter como um que recupera cerca de 300 vezes mais. Manter um agente é infraestrutura, utilização de modelos e as pessoas que o revêem quando mudam os documentos, os sistemas ou o modelo. Com os nossos números, um ano de manutenção custa entre 30 e 40 por cento do custo de construir; no software clássico, entre 10 e 20. Depende do modelo e do preço dos tokens, e vai mudar.

Construir paga-se uma vez. Manter paga-se para sempre.



Duas são sobre dados. Duas são decisões vossas.

Aceder e compreender resolvem-se com trabalho: permissões, formatos, significado acordado. Responder e compensar não se compram: alguém tem de decidir quem é o dono do agente, quais são as regras e se compensa. Um piloto contorna as duas primeiras com dados escolhidos a dedo e salta as duas últimas, porque há sempre alguém a vigiar e ninguém lhe pede ainda que se pague. É por isso que tantos pilotos funcionam e tão poucos chegam à produção.

As respostas constroem-se uma vez

As perguntas fazem-se a cada agente. As respostas constroem-se uma vez, numa plataforma de agentes com uma peça por pergunta. Identidade e acessos, para que o próximo agente receba o seu cartão como o dos contratos. Formatos legíveis por máquinas e uma camada semântica, para que o significado seja acordado uma vez e não uma vez por agente. Um registo com dono e regras aplicadas enquanto o agente corre, que é o que tinha o agente de inventário e não tinha o assistente de voz. Custo e valor por agente, manutenção incluída, que foi o que travou o agente dos relatórios de visita. Em tudo isto, cada execução fica registada e avaliada: o resultado, o custo e se um modelo novo faz melhor do que o anterior. Por baixo, segurança e conformidade, com uma parte comum e uma parte por agente; desenhámos essa camada com base no AI Act, na ISO/IEC 42001, no NIST AI RMF e no OWASP Top 10 for Agentic Applications, como referência, não como certificado. É uma primeira versão e vai mudar com a tecnologia. Mesmo assim, já suporta agentes de grande escala. A Mariana, a assistente de clientes do Continente, teve mais de 90.000 conversas completas em setembro. Foi construída internamente com o que aprendemos no programa, e cada conversa fica registada e auditável: que ferramentas usou, quanto tempo demorou e quanto custou.



O que eu faria na segunda-feira

Peguem no vosso piloto mais avançado e escrevam as quatro respostas, uma linha cada. Se alguma linha ficar em branco, é aí que ele vai parar quando sair do piloto. As duas primeiras são trabalho. As duas últimas precisam de quem manda sentado a decidir, e essa é a reunião mais difícil.

Estamos apenas na versão um. Vamos continuar a mudar e a melhorar o nosso processo e a nossa plataforma, e estamos disponíveis para discutir e aprender convosco. Não é preciso aprender isto sozinho.

Hélder Dias Ribeiro, Chief Digital & Technology Officer, Sonae MC. ·
www.helderdiasribeiro.net/4Questions/pt