

O controle rápido de riscos de fronteira

Abreviações e acrônimos

Nenhum conhecimento prévio é presumido. Cada acrônimo usado neste documento está escrito por extenso abaixo.

IA : inteligência artificial.

AGI : inteligência artificial geral, sistemas hipotéticos com capacidade cognitiva ampla comparável à humana.

O risco que muda de natureza

Um sistema que responde é uma ferramenta. Um sistema que age é um agente. Quando a IA começa a executar tarefas em cadeia, chamar outras ferramentas e tomar decisões intermediárias sem consulta, o risco muda de natureza. Dez perguntas para saber se você já cruzou essa linha.

As dez perguntas

N	Pergunta
1	Algum sistema executa ações (envia, compra, altera, publica) sem aprovação humana?
2	Algum sistema chama outros sistemas ou ferramentas por conta própria?
3	Existe um limite escrito do que cada agente pode fazer?
4	Existe um botão de parada que funciona em segundos, não em reuniões?
5	As ações do agente são registradas passo a passo?
6	Alguém revisa amostras das cadeias de ação do agente?
7	O agente pode gastar dinheiro ou comprometer a organização?
8	O agente interage com clientes ou cidadãos diretamente?
9	Você testou o que o agente faz diante de instruções ambíguas?
10	Existe um plano para o dia em que dois agentes interajam entre si?

Como ler o resultado

Cada resposta que revela ausência de controle (não nas perguntas 3, 4, 5, 6, 9, 10; sim nas perguntas 1, 2, 7, 8 sem os controles correspondentes) soma um ponto de exposição.

- **0 a 2** : fronteira governada.
- **3 a 5** : o agente corre à frente da governança. Feche as lacunas antes de ampliar.
- **6 ou mais** : você opera agentes sem guarda-corpos. Comece pelo botão de parada.

O próximo passo

O controle gratuito mede a exposição. A arquitetura completa de guarda-corpos, os padrões multiagente e a avaliação de prontidão vivem no Frontier and AGI Governance Kit da AI GOVERNANCE CHAIN™.

AI GOVERNANCE CHAIN™

Estrutura de governança de IA

© 2026 Mathieu K. Gouanou. Todos os direitos reservados.

Membro do Conselho Consultivo de Harvard Business Review.

Padrão do documento: governança de IA centrada no ser humano.