Quando mais IA deixa de significar melhor IA

O limite entre avanço tecnológico, eficiência real e risco institucional

Jandislei Antonio Genova

Resumo

Este artigo examina o limite plausível entre o avanço dos modelos de inteligência artificial, a eficiência real e o risco institucional. Sustenta-se que o progresso técnico não se converte automaticamente em valor institucional, especialmente quando modelos mais poderosos são integrados a processos frágeis, ambientes de dados pouco confiáveis ou estruturas de governança insuficientes. A partir da relação heurística V = B - C - R, o texto compreende o valor institucional da IA como o equilíbrio entre benefícios esperados, custos totais e riscos esperados. Discute-se eficiência marginal, custos materiais e energéticos, alucinação, confiança humana excessiva, governança sociotécnica e soberania digital. A tese central é que a fronteira da IA não será definida apenas pelos modelos mais poderosos, mas pela capacidade das instituições de converter potência técnica em uso responsável, auditável e socialmente legítimo.

Palavras-chave: inteligência artificial; eficiência institucional; governança de IA; risco institucional; soberania digital; influência cognitiva estrutural.

Abstract

This article examines the practical threshold at which advances in artificial intelligence models cease to generate proportional real-world efficiency and begin to increase institutional risk. It argues that technological progress does not automatically translate into institutional value, especially when more powerful models are integrated into fragile processes, unreliable data environments, or weak governance structures. Using the heuristic relation V = B - C - R, the article frames AI's institutional value as the balance between expected benefits, total costs, and expected risks. It discusses marginal efficiency, material and energy costs, hallucination, human overreliance, sociotechnical governance, and digital sovereignty. The central claim is that the frontier of AI will not be defined only by the most powerful models, but by the ability of institutions to convert technical capacity into accountable, auditable, and socially legitimate use.

Keywords: artificial intelligence; institutional efficiency; AI governance; institutional risk; digital sovereignty; structural cognitive influence.

Durante os últimos anos, a corrida da inteligência artificial foi apresentada como uma disputa por modelos cada vez maiores, mais rápidos, mais autônomos e mais inteligentes. Mais parâmetros, mais contexto, mais agentes, mais velocidade, mais capacidade multimodal, mais desempenho em testes e mais automação.

A narrativa dominante sugere que o próximo salto técnico resolverá o problema anterior: o modelo que hoje erra será superado por outro mais robusto; a alucinação será reduzida; a resposta será mais precisa; o raciocínio será mais robusto; a ferramenta será mais integrada; o agente será mais autônomo; a produtividade aumentará.

Essa narrativa tem parte de verdade. A IA continuará avançando, e modelos melhores serão incorporados a empresas, governos, escolas, hospitais, escritórios, bancos e serviços públicos. Mas talvez esse debate esteja olhando para a inteligência artificial em altitude errada.

A pergunta decisiva não é apenas até onde os modelos podem avançar. A pergunta mais difícil é outra: em que ponto o avanço dos modelos deixa de produzir valor real proporcional e passa a ampliar custos, opacidade, dependência e risco?

O ponto é preciso: avanço técnico não equivale automaticamente a eficiência real. Um modelo pode ser mais poderoso e, ainda assim, produzir pouco valor se for inserido em um processo ruim. Pode responder melhor, mas operar sobre dados frágeis. Pode automatizar tarefas, mas acelerar decisões mal estruturadas. Pode reduzir tempo, mas aumentar dependência. Pode impressionar no teste, mas fracassar no contexto institucional em que será usado.

É aqui que surge o limite invisível da inteligência artificial: não um limite absoluto da tecnologia, mas um limite de conversão entre capacidade técnica e valor real.

Em termos heurísticos, o valor institucional líquido da IA pode ser pensado pela seguinte relação:

$$\mathbf{V = B - C - R}$$

Nessa formulação, B representa os benefícios esperados: produtividade, qualidade, precisão, velocidade, escala e apoio à decisão. C representa os custos totais: energia, infraestrutura, computação, treinamento, integração, manutenção, governança, profissionais qualificados e dependência tecnológica. R representa o risco esperado, ponderado por probabilidade, gravidade, reversibilidade e exposição institucional: erro, alucinação, vazamento de dados, discriminação, falha de supervisão, dano jurídico, reputacional, social ou institucional.

Para fins de avaliação concreta, essas variáveis devem ser normalizadas ou ponderadas conforme o contexto institucional, a finalidade do sistema e a gravidade dos impactos possíveis.

Essa fórmula não pretende reduzir a IA a uma conta linear. Ela funciona como uma salvaguarda conceitual. Sob o critério estrito de valor institucional líquido, a expansão deixa de produzir ganho marginal positivo quando $\Delta V < 0$, isto é, quando o ganho marginal de benefícios se torna inferior ao aumento marginal de custos e riscos. Em outras palavras:

$$\mathbf{\Delta B < \Delta C + \Delta R}$$

Esse talvez seja o ponto que o debate público sobre IA ainda evita enfrentar. Nem todo avanço de desempenho justifica o custo de sua implementação. Nem todo ganho de precisão compensa o aumento de opacidade. Nem toda automação reduz risco. Nem todo modelo mais sofisticado produz decisões melhores. Nem toda redução de tempo significa aumento de qualidade.

A inteligência artificial pode evoluir tecnicamente e, ao mesmo tempo, produzir eficiência institucional decrescente. Esse fenômeno não é estranho à economia: tecnologias maduras frequentemente encontram pontos de rendimento marginal decrescente, nos quais cada novo salto exige mais investimento para gerar ganhos proporcionalmente menores.

Na IA, esse problema é ainda mais sensível porque o custo não é apenas financeiro. É energético, computacional, ambiental, organizacional, jurídico, cognitivo e institucional. Em geral, modelos maiores ou usos realizados em maior escala tendem a exigir mais capacidade computacional, data centers, chips, energia, água, resfriamento, conectividade, equipes especializadas, contratos complexos, segurança cibernética e integração com sistemas existentes.

A literatura técnica também mostra que desempenho não depende simplesmente de aumentar parâmetros. Estudos sobre compute-optimal scaling indicam que, sob determinado orçamento computacional, tamanho do modelo e volume de dados precisam ser equilibrados; em alguns casos, modelos menores e mais bem treinados podem superar modelos muito maiores (HOFFMANN et al., 2022).

O custo material da IA também não está restrito ao treinamento. Pesquisas sobre inferência mostram que sistemas generativos de propósito geral podem ser energeticamente muito mais caros do que sistemas específicos para tarefas delimitadas, o que exige ponderar utilidade, escala e impacto ambiental antes da adoção indiscriminada (LUCCIONI; JERNITE; STRUBELL, 2024). No mesmo sentido, a Agência Internacional de Energia projeta forte crescimento do consumo elétrico de data centers até 2030, impulsionado pela expansão da IA e da infraestrutura digital (INTERNATIONAL ENERGY AGENCY, 2025).

A pergunta, portanto, não é apenas se o modelo é melhor. É: melhor para quem, em qual contexto, a que custo e sob qual risco?

Uma IA avançada em um hospital sem dados confiáveis, sem integração clínica, sem governança e sem revisão qualificada pode apenas acelerar a confusão. Uma IA avançada em um escritório jurídico sem protocolo de verificação pode produzir textos mais rápidos, mas aumentar o risco de erro técnico. Uma IA avançada em banco, RH ou administração pública pode sofisticar exclusões se os critérios, dados e objetivos forem frágeis.

A potência do modelo não corrige sozinha a fragilidade da instituição. Às vezes, apenas acelera seus efeitos.

Esse é um ponto central para qualquer discussão séria sobre IA aplicada. O valor da inteligência artificial não está apenas no modelo; está na arquitetura de uso. Modelo, por si só, não é solução. Modelo é componente.

O valor real surge da combinação entre modelo, dados, finalidade, contexto, governança, supervisão humana, métricas adequadas, segurança, auditabilidade e responsabilidade institucional. Essa conclusão dialoga com pesquisas empíricas sobre IA generativa no trabalho: ganhos de produtividade existem, mas dependem do contexto operacional, do tipo de tarefa, da integração ao fluxo de trabalho e do perfil dos usuários (BRYNJOLFSSON; LI; RAYMOND, 2025).

Essa é a razão pela qual marcos de gestão de risco, como o AI Risk Management Framework do NIST, tratam a IA como sistema sociotécnico: os riscos não emergem apenas do modelo ou dos dados, mas da forma como pessoas, organizações e instituições projetam, implantam, supervisionam e utilizam esses sistemas (NIST, 2023).

Sem essa arquitetura, a IA pode produzir uma modernização aparente: a organização passa a usar tecnologia de ponta, mas continua decidindo mal, medindo mal, documentando mal, supervisionando mal e respondendo mal pelos próprios efeitos.

Há também outro limite, menos confortável: o limite do erro. Parte do discurso tecnológico sugere que modelos maiores tenderão a eliminar falhas relevantes. Mas, no mundo real, nem todo erro decorre apenas de insuficiência computacional. Há erros que nascem da própria natureza dos dados e dos contextos.

Dados podem ser incompletos, enviesados, contraditórios ou desatualizados. Podem não representar quem ficou fora da base. Podem capturar apenas os sobreviventes do sistema. Podem transformar desigualdades anteriores em padrões futuros.

Além disso, há áreas em que a resposta correta não é puramente estatística. Direito, saúde, educação, políticas públicas, gestão institucional e relações humanas envolvem ambiguidade, valor, contexto, prudência, responsabilidade e interpretação. Nesses campos, a IA pode apoiar, organizar, sugerir, comparar e acelerar. Mas não elimina a necessidade de julgamento.

O problema começa quando a aproximação probabilística passa a ser confundida com compreensão; quando fluência passa a ser confundida com verdade; quando uma resposta bem escrita passa a ser tomada como resposta correta; quando uma métrica elevada passa a ser lida como garantia de sentido.

Essa preocupação não é meramente abstrata. Estudos recentes sobre alucinações em modelos de linguagem indicam que respostas plausíveis, porém incorretas, persistem mesmo em sistemas avançados, especialmente quando procedimentos de treinamento e avaliação recompensam o palpite confiante em vez da abstenção diante da incerteza (KALAI et al., 2025).

Esse é o limite cognitivo da IA. Quanto melhor ela parece, mais perigosa pode se tornar a confiança indevida. Uma ferramenta claramente limitada convida à cautela. Uma ferramenta sofisticada, fluente, rápida e convincente pode reduzir a vigilância humana justamente porque parece segura.

O erro não desaparece. Ele muda de aparência. Passa a circular com linguagem técnica, estrutura lógica, tom confiante e autoridade aparente.

Esse ponto é decisivo para o que denomino influência cognitiva estrutural. A IA não influencia apenas quando entrega uma resposta errada. Ela influencia quando reorganiza o ambiente de confiança: o que parece relevante, o que parece correto, o que parece suficiente, o que parece revisado, o que parece tecnicamente validado.

O risco não está apenas no erro da máquina. Está na reorganização da confiança humana diante da máquina.

Por isso, a fronteira real da IA talvez não esteja apenas nos laboratórios que produzem modelos mais capazes. Está nas instituições que precisarão decidir como, quando, por que e sob quais limites esses modelos serão usados.

A pergunta 'qual é o melhor modelo?' é insuficiente. A pergunta correta é mais difícil: qual é a melhor arquitetura institucional para usar esse modelo sem transformar capacidade técnica em risco sistêmico?

Essa pergunta exige maturidade. Exige reconhecer que IA não é mágica. Exige abandonar a ideia de que toda adoção tecnológica é progresso. Exige distinguir produtividade de qualidade, velocidade de julgamento, automação de responsabilidade, escala de justiça, precisão estatística de adequação contextual e sofisticação técnica de governança real.

A corrida por modelos mais inteligentes pode gerar uma ilusão: a de que o problema está sempre no modelo anterior e a solução sempre no próximo modelo. Mas muitas falhas não decorrem apenas de falta de inteligência artificial. Decorrem de falta de inteligência institucional.

Organizações que não sabem definir finalidade usarão IA para acelerar ambiguidades. Organizações que não têm dados confiáveis usarão IA para produzir conclusões frágeis com aparência sofisticada. Organizações sem cultura de revisão usarão IA para terceirizar responsabilidade. Estados sem capacidade regulatória usarão IA sem conseguir auditar seus efeitos. Sociedades com baixo letramento digital serão expostas a sistemas que não compreendem e não conseguem contestar.

Nesse cenário, mais IA pode não significar melhor IA. Pode significar mais dependência, mais opacidade, mais custo, mais assimetria, mais dificuldade de responsabilização e mais distância entre quem controla a infraestrutura e quem sofre seus efeitos.

A preocupação também aparece em estratégias recentes de IA e cibersegurança. Ao associar capacidades avançadas, infraestrutura de avaliação, acesso seguro a modelos e resiliência digital, a Comissão Europeia evidencia que a IA de fronteira já não pode ser tratada apenas como ferramenta de produtividade, mas como componente de soberania tecnológica e segurança institucional (EUROPEAN COMMISSION, 2026).

Esse é também um eixo atualmente em desenvolvimento em minha pesquisa sobre soberania digital e novas formas de dependência tecnológica. A soberania capturada não aparece apenas quando um país depende de modelos estrangeiros, nuvens privadas, chips importados ou data centers externos. Ela também aparece quando instituições e pessoas passam a depender de sistemas que não conseguem compreender, auditar, substituir ou governar.

A dependência pode ser técnica, mas também pode ser cognitiva. A organização pode começar a perder capacidade decisória própria quando transfere à ferramenta não apenas tarefas, mas critérios de relevância, prioridade e suficiência. O profissional pode perder tolerância ao esforço de formulação própria. O estudante pode deixar de tolerar a dúvida. O usuário pode deixar de buscar compreensão e passar a buscar apenas resposta. O Estado pode digitalizar processos sem garantir capacidade social de contestação.

A sociedade passa a operar dentro de uma infraestrutura que promete eficiência, mas cobra, em troca, parcelas de autonomia.

Isso não significa rejeitar a IA. Significa recolocá-la em sua devida posição: infraestrutura de apoio, não substituto invisível do julgamento; instrumento de ampliação de capacidades, não

justificativa para enfraquecimento da responsabilidade; tecnologia poderosa, mas não autoridade cognitiva automática.

O limite plausível da IA, portanto, não é o limite técnico da inteligência artificial. É o limite econômico, energético, jurídico, cognitivo e institucional da sua conversão em valor real. Esse limite aparece quando o custo de usar modelos mais sofisticados supera o benefício concreto; quando o risco de erro permanece, mas a confiança humana aumenta; quando a governança não acompanha a automação; quando a infraestrutura se concentra em poucos fornecedores; quando o acesso pleno passa a ser seletivo; quando o uso da IA amplia desigualdades em vez de reduzi-las.

A tecnologia continuará avançando. Mas a pergunta pública precisa amadurecer. Não basta perguntar qual será o próximo modelo. É preciso perguntar qual será o próximo custo, a próxima dependência, a próxima assimetria, o próximo risco e a próxima responsabilidade deslocada.

Talvez a maior ingenuidade contemporânea seja imaginar que toda inteligência adicional do modelo se converte automaticamente em inteligência social. Não se converte. Entre a capacidade técnica e o valor real existe uma ponte. Essa ponte se chama governança.

Sem ela, a IA pode continuar avançando, mas as instituições podem continuar errando, apenas com mais velocidade, mais escala e mais aparência de precisão. O futuro da inteligência artificial não será decidido apenas por quem construir os modelos mais poderosos. Será decidido por quem souber reconhecer quando acréscimos de capacidade deixam de produzir benefício proporcional e passam a ampliar custo, opacidade, dependência e risco.

Esse é o limite invisível da IA: não o limite da máquina, mas o limite da nossa capacidade de transformar potência técnica em responsabilidade humana.

Referências selecionadas

BENDER, Emily M.; GEBRU, Timnit; McMILLAN-MAJOR, Angelina; SHMITCHELL, Shmargaret. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, p. 610-623, 2021. DOI: 10.1145/3442188.3445922.

BRYNJOLFSSON, Erik; LI, Danielle; RAYMOND, Lindsey R. Generative AI at Work. The Quarterly Journal of Economics, v. 140, n. 2, p. 889-942, 2025. DOI: 10.1093/qje/qjae044.

EUROPEAN COMMISSION. EU Action Plan on Cybersecurity and Artificial Intelligence. COM(2026) 577 final. Strasbourg, 7 July 2026.

HOFFMANN, Jordan et al. Training Compute-Optimal Large Language Models. Advances in Neural Information Processing Systems, 2022. Disponível em: arXiv:2203.15556.

INTERNATIONAL ENERGY AGENCY. Energy and AI. Paris: IEA, 2025.

KALAI, Adam Tauman; NACHUM, Ofir; VEMPALA, Santosh S.; ZHANG, Edwin. Why Language Models Hallucinate. arXiv, 2025. Disponível em: arXiv:2509.04664.

LUCCIONI, Alexandra Sasha; JERNITE, Yacine; STRUBELL, Emma. Power Hungry Processing: Watts Driving the Cost of AI Deployment? Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency, 2024. DOI: 10.1145/3630106.3658542.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. Artificial Intelligence Risk Management Framework (AI RMF 1.0). Gaithersburg: NIST, 2023. DOI: 10.6028/NIST.AI.100-1.

Sobre o autor

Jandislei Antonio Genova é advogado, pesquisador independente, com atuação no serviço público de saúde em ambiente hospitalar de alta complexidade. Pós-graduado em Gestão em Saúde, desenvolve pesquisas sobre inteligência artificial, autonomia decisória, governança digital, vulnerabilidade, responsabilidade institucional e soberania tecnológica. É autor de obras sobre inteligência artificial, plataformas digitais e arquitetura jurídica da decisão.