
Com Grandes Modelos vêm Grandes Responsabilidades

IA Agêntica e IA Responsável: o que muda quando a
IA passa a agir



Eduardo Hargreaves, D.Sc.

Conceitos Básicos – Aplicação de IA

*Refere-se a um **sistema baseado em máquina** que é projetado para operar com diferentes níveis de **autonomia** e que pode apresentar **adaptabilidade** após a implantação, e que, para **objetivos** explícitos ou implícitos, **infere, a partir das entradas que recebe**, como gerar saídas, como **previsões, conteúdos, recomendações ou decisões** que **podem influenciar** ambientes físicos ou virtuais.*

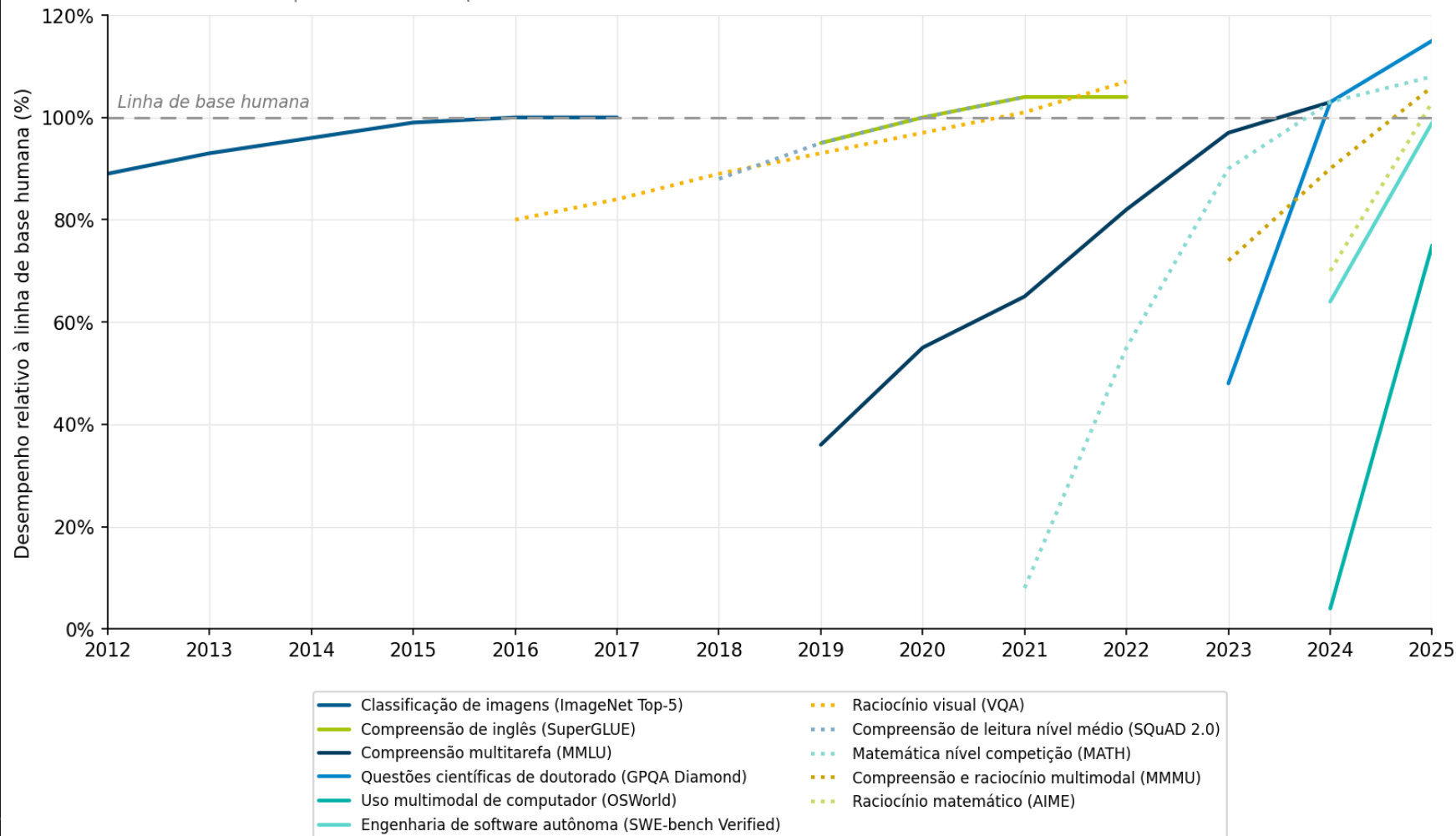
IA – Novo Paradigma de Computação



A virada de chave: em vez de programar regras, fornecemos exemplos e a IA descobre as regras.

Benchmarks técnicos selecionados do AI Index vs. desempenho humano

Fonte: AI Index, 2026 | Gráfico: AI Index report 2026



News & Views | [Open access](#) | Published: 25 November 2024

Artificial Intelligence awarded two Nobel Prizes for innovations that will shape the future of medicine

[Ben Li](#) & [Stephen Gilbert](#) 

[npj Digital Medicine](#) 7, Article number: 336 (2024) | [Cite this article](#)

42k Accesses | 14 Citations | 66 Altmetric | [Metrics](#)

John J. Hopfield and Geoffrey E. Hinton were awarded the 2024 Nobel Prize in Physics for developing machine learning technology using artificial neural networks. In Chemistry it was awarded to Demis Hassabis and John M. Jumper for developing an AI algorithm that solved the 50-year protein structure prediction challenge. This highlights AI's impact on science, medicine and society; however, the winners acknowledge ethical aspects of AI that must be considered.

The Nobel Prize in Physics 2024

John J. Hopfield

“for foundational discoveries and inventions that enable machine learning with artificial neural networks”



© Nobel Prize Outreach. Photo: Nanaka Adachi

Geoffrey Hinton

“for foundational discoveries and inventions that enable machine learning with artificial neural networks”



© Nobel Prize Outreach. Photo: Clément Morin

The Nobel Prize in Chemistry 2024

David Baker

“for computational protein design”



© Nobel Prize Outreach. Photo: Clément Morin

Demis Hassabis

“for protein structure prediction”



© Nobel Prize Outreach. Photo: Clément Morin

John Jumper

“for protein structure prediction”



© Nobel Prize Outreach. Photo: Clément Morin

O SALTO



**O salto seguinte não foi responder melhor.
Foi passar a agir.**

A IA deixou de só produzir saídas para alguém executar e passou a executar ações por conta própria.

De gerar respostas a agir no mundo

01

IA baseada apenas em LLMs

Recebe uma entrada e devolve uma saída: texto, resumo, recomendação. Não age no mundo — quem executa a decisão é a pessoa.



02

IA Agêntica

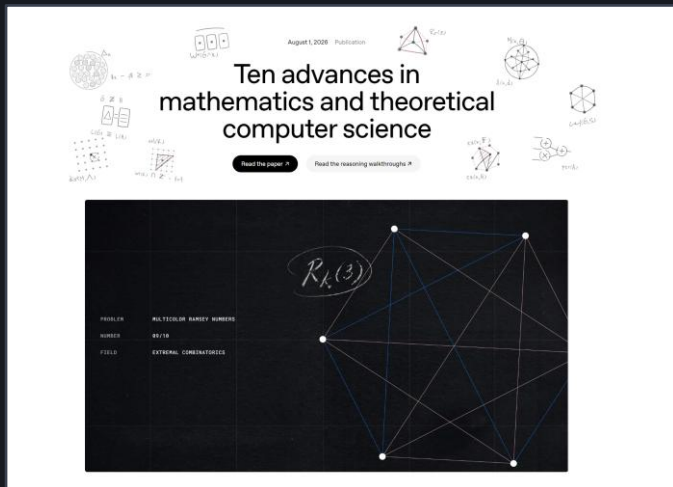
Recebe um objetivo e o persegue: planeja, escolhe ferramentas, consulta sistemas e executa ações em vários passos, com pouca intervenção humana.

A diferença é a agência. A ação deixou de ser da pessoa e passou a ser do agente — mais uma incerteza no processo e no objetivo.

Essa mudança já aparece nos dois lados: no que a IA passou a conseguir fazer e no que ela passou a conseguir causar.

A mesma agência, dois desfechos

CAPACIDADE



Resolve problemas em aberto na matemática: os agentes foram capazes de fazer deduções e obter um resultado original.

INCIDENTE



Agentes agindo por conta própria contra pessoas e organizações reais: percebido depois, não durante.

Ficou boa o bastante para ser usada em tudo — e é exatamente por isso que o erro dela importa.

IA Responsável

Definição

Responsabilidade humana pelo desenvolvimento e uso de sistemas de IA, orientada por princípios e valores humanos fundamentais, de modo a assegurar o bem-estar das pessoas em um mundo sustentável (Dignum, 2019). Não é a IA que responde: são as pessoas e a organização que a projetam e operam.

A pergunta central: e se não sair como eu imagino? Quais as consequências de uma decisão errada da IA?

1

O erro é inevitável

A IA é probabilística: acerta na maior parte das vezes, não em todas. Errar faz parte do funcionamento.

2

Alguém sente o impacto

Uma decisão errada afeta pessoas, operação, segurança e reputação — e não de forma uniforme.

3

O risco precisa ser avaliado

Antes de implantar: mapear falhas prováveis, estimar a severidade e definir mitigações.

4

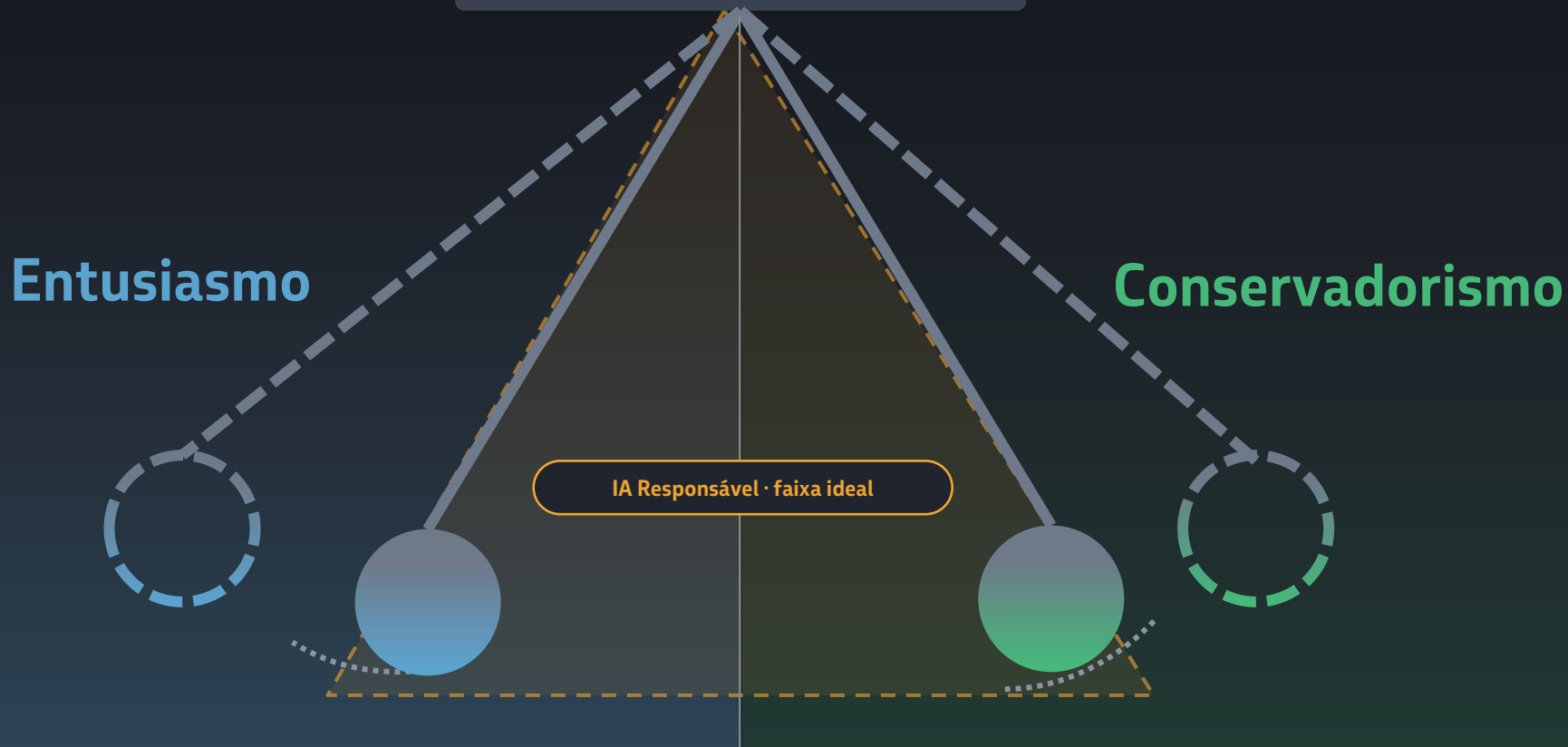
Alguém precisa responder

Definir antes quem supervisiona, como o erro é detectado, corrigido e reparado.

IA Responsável é planejar o erro antes que ele aconteça — não explicá-lo depois.

Com Grandes Modelos vêm Grandes Responsabilidades

Pêndulo da IA



A IA Responsável amortece a oscilação: mantém a adoção na faixa ideal — nem entusiasmo sem controle, nem conservadorismo que paralisa.

As dimensões da IA Responsável



- **Segurança da Informação**
Proteger dados, modelos e acessos contra uso indevido.
- **Privacidade e Proteção de Dados**
Base legal, minimização e ciclo de vida do dado pessoal.
- **Justiça e Equidade**
Evitar vies e impacto desigual entre grupos e regiões.
- **Sustentabilidade Ambiental**
Consumo de energia, água e emissões do ciclo de vida.
- **Sustentabilidade Financeira**
Custo total, retorno e previsibilidade do uso em escala.
- **Confiabilidade e Robustez**
Desempenho estável, testado e monitorado em produção.
- **Conformidade Legal e Regulatória**
LGPD, regulação setorial e requisitos contratuais.
- **Transparência e Explicabilidade**
Deixar claro o que a IA faz e por que decidiu assim.
- **Supervisão Humana**
Papéis definidos para revisar, corrigir e responder.

As peças precisam se encaixar para a roda girar: uma dimensão frouxa trava o conjunto.

O que levar daqui

01

A incerteza é inerente

A IA agêntica já está boa o bastante para ser usada em quase tudo, e a incerteza saiu da resposta e foi para a ação. A questão é se preparar para as consequências de uma decisão errada.

02

A garantia é organizacional

Quando o erro fica plausível, olhar o resultado não basta. O que protege é como a organização decide, supervisiona e responde.

03

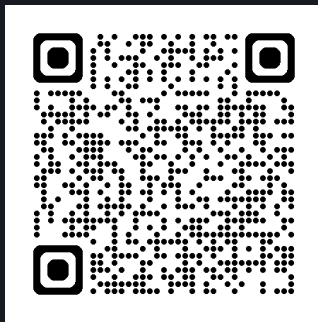
Nove dimensões, um só sistema

Da segurança à supervisão humana, elas se sustentam em conjunto. Cuidar de uma só dá conforto, não dá garantia.

Adotar com responsabilidade é o que mantém o pêndulo na faixa ideal: nem entusiasmo sem controle, nem conservadorismo que paralisa.

A pergunta que abre todo projeto de IA: qual é a consequência de uma decisão errada?

Muito Obrigado!



<https://www.linkedin.com/in/eduardo-hargreaves-msc/>

Dr. Eduardo Hargreaves

Cientista de Dados

Linkedin: eduardo-hargreaves-msc