

ANTI/HYPE · FORESIGHT

FULL REPORT

A Reprecificação da Cognição

Inteligência Artificial em 2030, 2040 e 2050 | Full Foresight Report

Gui Zanoni · ANTI/HYPE

Julho 2026

01

Sumário executivo

A IA não é uma tecnologia que a sua organização adota. É uma reprecificação do custo de pensar, e tudo que era caro porque cognição era cara vai ser reprecificado junto: trabalho, software, educação, decisão, poder. O veredito deste report é direto: a direção da mudança é quase certa, o ritmo é a incerteza que vale dinheiro, e a distribuição do poder é a incerteza que vale tudo o resto.

53%

da população adulta em grandes mercados adotou IA generativa em 3 anos, mais rápido que PC e internet

Fonte: Stanford AI Index 2026

US\$ 5,5 tri

de capex global em IA projetado até 2030

Fonte: JPMorgan Research 2026

945 TWh

consumo elétrico de data centers em 2030, o dobro de 2024 e mais que o Japão inteiro

Fonte: IEA, Energy and AI 2025

95%

dos pilotos corporativos de IA generativa sem impacto mensurável no resultado

Fonte: MIT NANDA 2025

Cinco achados sustentam o veredito. Primeiro: a capacidade técnica está correndo mais rápido que a capacidade institucional de absorver, e essa é a tensão que define a década. Agentes de IA saltaram de 12% para 66% de sucesso em tarefas reais de computador em um ano (OSWorld, Stanford AI Index 2026), mas 88% dos pilotos corporativos de agentes nunca chegam à produção (Anaconda e Forrester 2026). O problema de 2026 a 2030 não é o modelo, é a organização.

Segundo: a economia da IA virou uma aposta de infraestrutura sem precedente. Os cinco maiores hyperscalers devem gastar cerca de US\$ 700 bilhões em capex só em 2026, quase o dobro de 2025, com projeção de ultrapassar US\$ 1 trilhão em 2027 (JPMorgan e Bank of America, 2026). Uma correção financeira no caminho é provável e não muda a direção estrutural: depois do estouro da bolha pontocom, a fibra ótica ficou no chão e virou a internet que usamos.

Terceiro: energia virou o limitador físico da inteligência. O consumo elétrico de data centers dobra até 2030 e a fatia de IA triplica (IEA 2026). Pela primeira vez desde a eletrificação, a fronteira de uma tecnologia de informação esbarra em turbina, transformador e linha de transmissão. Quem controla energia passa a controlar cognição.

Quarto: o trabalho está sendo reorganizado antes de ser substituído. Nos EUA, a IA foi o motivo líder de demissões por três meses consecutivos em 2026 (Challenger, Gray & Christmas), mas a projeção de 502 mil cortes ligados a IA em 2026 representa 0,4% da força de trabalho americana (Gazeta Mercantil sobre dados Challenger, 2026), e 55% dos empregadores dizem se arrepender de demissões motivadas por IA (Forrester 2026). O padrão histórico se repete: a tarefa morre antes do cargo, o cargo muda antes de sumir.

Quinto: o mundo de 2040 não será decidido pela pergunta "a IA vai funcionar", e sim por duas incertezas independentes: se a capacidade continua saltando ou estabiliza num platô sofisticado, e se o poder computacional se concentra em poucos atores ou se difunde. O cruzamento dessas duas forças gera os quatro mundos possíveis do médio prazo, e é neles que mora a decisão de capital.

A janela de decisão: 2026 a 2029. É o intervalo em que redesenhar processos em torno de IA ainda é vantagem competitiva e não pré-requisito de sobrevivência. Empresas que redesenham fluxos antes de

escolher ferramentas têm 2x mais chance de retorno financeiro significativo (McKinsey 2025). Depois de 2030, o custo de entrada não é a tecnologia, é a distância acumulada.



Quando pensar fica barato, o que fica caro é decidir.

VEREDITO CENTRAL

02

Introdução e framing

Questão focal: como será a inteligência artificial, e o mundo reorganizado por ela, em 2030, 2040 e 2050, e quais as implicações estratégicas para organizações e profissionais que precisam decidir agora?

Escopo do estudo:

- Sujeito: IA como sistema sociotécnico, incluindo modelos, agentes, robótica, infraestrutura de computação e energia, e os comportamentos e instituições ao redor.
- Horizontes: definidos por solicitação do cliente em três marcos, 2030 (H1, curto), 2040 (H2, médio) e 2050 (H3, longo), com exercício distópico separado no teto de 50 anos. O ano-alvo é a quantidade de mudança permitida, não uma data literal.
- Recorte geográfico: global, com lente adicional sobre o Brasil, onde a adoção corre à frente da governança.
- Entrega: cenários acionáveis, implicações por ator e framework de monitoramento.

Contexto de partida, as três tensões estruturais do presente. Tensão um: capacidade contra confiabilidade. Os modelos de fronteira ganharam medalha de ouro na Olimpíada de Matemática, mas o melhor modelo lê relógio analógico corretamente só 50,1% das vezes (Stanford AI Index 2026). O salto de benchmark não é salto de produção: pesquisas de 2026 mediram uma lacuna de 37% entre desempenho de laboratório e desempenho real de agentes em ambiente corporativo (Kili Technology 2026, sobre estudos revisados). Tensão dois: investimento contra retorno. Gartner projeta US\$ 2,53 trilhões de gasto global em IA em 2026, enquanto o MIT mediu que 95% dos pilotos de IA generativa não produziram impacto mensurável no resultado. O dinheiro corre na frente da prova. Tensão três: velocidade privada contra velocidade pública. A União Europeia adiou em 16 meses as obrigações de alto risco do AI Act (Digital Omnibus, acordo de maio de 2026, Conselho da UE) porque o mercado não conseguia cumprir o cronograma que a própria lei criou. O Brasil discute o PL 2338 desde 2023 e a votação na Câmara ficou para 2026 em ano eleitoral. Reguladores estão legislando sobre um alvo que se move mais rápido que o rito legislativo.

03

Metodologia

Este report roda, nesta ordem e de fato: varredura ambiental STEEP com dimensão de valores (STEEPV), radar de sinais com classificação de confiança, horizonte e tipo, mapeamento de drivers em grade de impacto por incerteza com seleção de eixos, Causal Layered Analysis em quatro camadas, Três Horizontes como espinha temporal, construção de cenários com matriz 2x2 de Schwartz no médio prazo e cenários ancorados por horizonte, implicações robustas e contingentes por ator do sistema, e framework de monitoramento por signposts.

Disciplina de fonte aplicada: triangulação para afirmações estruturais, rótulo de tier por fonte (primária, institucional, fornecedor, imprensa), ressalva explícita em projeções de mercado (firmas de projeção inflam, e o intervalo entre elas é o dado honesto) e nível de confiança declarado onde a evidência é fina. Projeções de PIB e de mercado aparecem sempre como intervalo entre fontes divergentes, nunca como número único.

Limite honesto do estudo: não há pesquisa primária de campo nem acesso a dado proprietário de laboratórios. Sobre trajetória de capacidade além de 2030, a evidência é estruturalmente fina para todo mundo, inclusive para quem constrói os modelos. Onde isso pesa, o texto sinaliza.

04

Scanning ambiental

Social: a adoção mais rápida da história, e a mais íntima

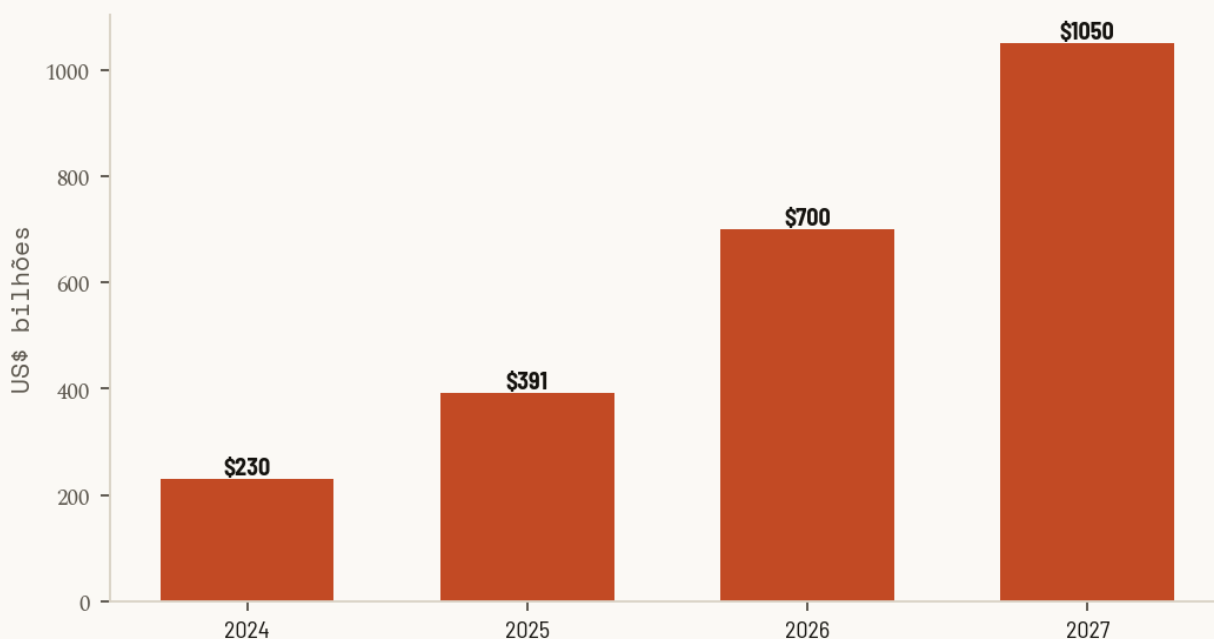
A IA generativa atingiu 53% de adoção populacional em grandes mercados em três anos, ritmo superior ao do PC e da internet (Stanford AI Index 2026). O que esse número não conta é a natureza do uso. Nos EUA, 72% dos adolescentes já usaram IA para companhia, e mais da metade desses usa de forma recorrente (Common Sense Media 2025). A OpenAI reportou cerca de 800 milhões de usuários semanais, dos quais cerca de 1 milhão por semana fala sobre suicídio com o sistema (dados citados pela APA, 2026). Isso é sinal forte de uma mudança de camada: a IA saiu da categoria ferramenta e entrou na categoria relação. Estudo conjunto OpenAI e MIT Media Lab encontrou que uso intenso diário de chatbot correlaciona com aumento de solidão (Phang et al., 2025). Driver de fundo: a economia da atenção migrando para a economia da intimidade sintética. No Brasil, 82% das organizações já liberam IA para pelo menos 20% dos profissionais (Deloitte, State of AI in the Enterprise 2026) e 47,4% dos profissionais usam ferramentas sem aprovação oficial, o Shadow AI (Abiacom 2026). O comportamento corre décadas à frente da política de uso.

Tecnológico: do chatbot ao trabalhador, com um teto visível de confiabilidade

2026 é o ano em que os modelos deixaram de ser interfaces de conversa e viraram executores. Praticamente todo lançamento de fronteira do primeiro semestre incluiu capacidade agêntica: tarefas longas de múltiplos passos, uso de computador, enxames de agentes (TeamDay 2026). Agentes saltaram de 12% para 66% de sucesso no OSWorld em um ano, a poucos pontos do desempenho humano em tarefas digitais estruturadas (Stanford AI Index 2026). No Humanity's Last Exam, benchmark montado por especialistas para ser difícil, os melhores modelos foram de 8,8% para cerca de 38% de acerto em pouco mais de um ano (IEEE Spectrum sobre AI Index 2026). Ao mesmo tempo, benchmarks clássicos saturaram (todo modelo de fronteira acima de 88% no MMLU) e perderam poder de diferenciação (Kili Technology 2026). O sinal técnico mais barulhento do período veio da China: o GLM-5 foi treinado inteiramente em chips Huawei Ascend, zero hardware americano (TeamDay 2026), e o DeepSeek V4 entregou desempenho quase de fronteira em código aberto a fração do custo. O custo por unidade de inteligência está em colapso: modelos intermediários a US\$ 3 por milhão de tokens empatando com flagships de US\$ 15. No mundo físico, o contraste é brutal: robôs têm 12% de sucesso em tarefas domésticas reais contra 89% em simulação (Stanford AI Index 2026). A Unitree embarcou mais de 5.500 humanoides em 2025 com preço médio do setor caindo de US\$ 85 mil em 2023 para US\$ 25 mil em 2025 (KraneShares 2026, sobre dados da empresa, marcar como fornecedor). Sinal recente forte; a curva de custo é o driver.

Econômico: a maior aposta de capital da história corporativa, com retorno ainda em disputa

Capex anual dos maiores hyperscalers

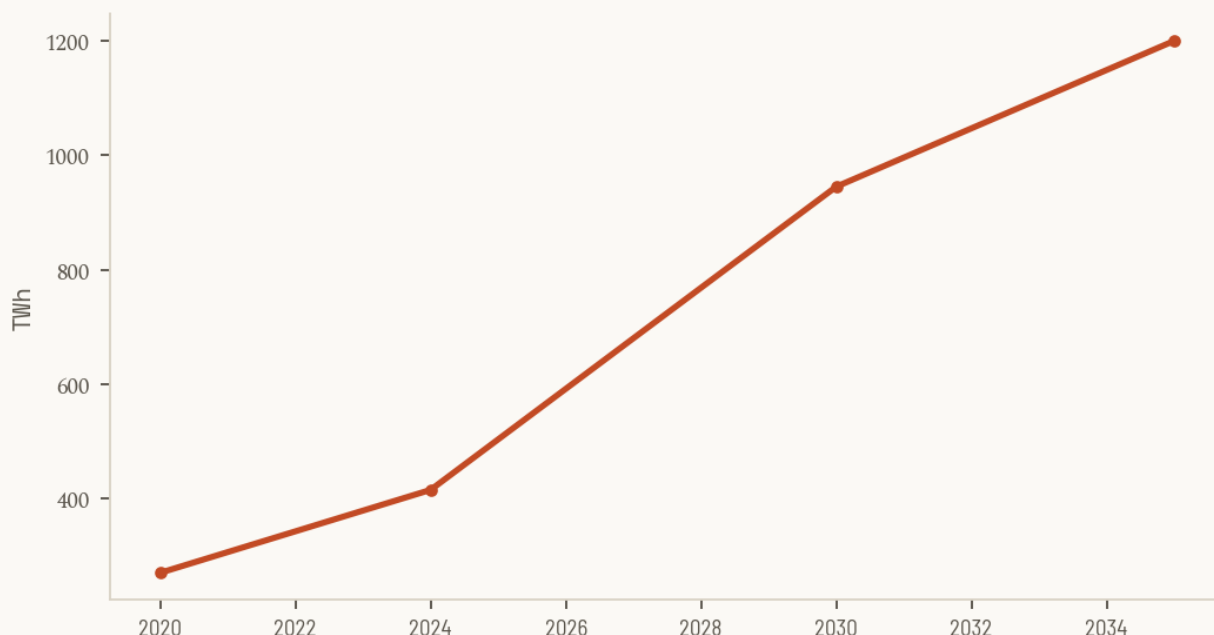


Fonte: Bloomberg via Guinness GI (2025), JPMorgan e Bank of America (2026); 2026 e 2027 são projeções de bancos, tratar como estimativa. O que significa: o gasto quase dobra por ano, e a receita ainda não acompanha.

Capex projetado dos hyperscalers: cerca de US\$ 700 bilhões em 2026, mais de US\$ 1 trilhão em 2027, US\$ 5,5 trilhões acumulados até 2030 (JPMorgan 2026, projeção de banco, tratar como estimativa interessada em cenário construtivo). Do outro lado da mesa: MIT mediu 95% de pilotos sem impacto em P&L, McKinsey mediu 88% de adoção com só 39% de impacto em EBIT, e a Gartner projeta mais de 40% dos projetos de IA agêntica cancelados até 2027. A assimetria mais citada do período: a OpenAI gastando na casa de US\$ 60 bilhões por ano em computação contra cerca de US\$ 13 bilhões de receita (estimativas de imprensa e análise, 2026, baixa verificabilidade, tratar com cautela). Sobre o impacto macroeconômico, o intervalo entre as casas é o dado: PwC projeta até US\$ 15,7 trilhões de contribuição ao PIB global até 2030 (fornecedor de projeção, historicamente inflado), Goldman Sachs fala em +7% de PIB global em dez anos, McKinsey em US\$ 2,6 a 4,4 trilhões por ano, e Acemoglu, no MIT, em torno de 1% em dez anos. Uma ordem de grandeza separa o teto do piso. Penn Wharton, no meio do caminho, projeta +1,5% de PIB até 2035 com pico de contribuição no início dos anos 2030. Leitura prática: ninguém sabe o tamanho, todo mundo concorda com o sinal positivo e com a concentração de ganho em quem redesenha processo, não em quem compra licença.

Ambiental: energia é o novo silício

Eletricidade consumida por data centers no mundo



Fonte: IEA, Energy and AI (2025). O que significa: em 2030, os data centers consomem mais que o Japão inteiro, e a fatia de IA triplica no período.

Data centers consumiram 415 TWh em 2024 e vão a cerca de 945 TWh em 2030, quase 3% da eletricidade global, com a fatia de IA triplicando no período e chegando a 1.200 TWh em 2035 (IEA, Energy and AI, 2025 e atualização 2026). Nos EUA, data centers vão consumir mais eletricidade que toda a indústria energointensiva somada, alumínio, aço, cimento e química, até 2030 (IEA 2025). O consumo de data centers de IA cresceu 50% só em 2025 (IEA 2026). A densidade de potência dos servidores de IA subiu 11 vezes entre 2020 e 2025 e deve quadruplicar de novo até 2027 (IEA 2026): um rack do tamanho de uma geladeira com demanda de pico de 65 residências. Consequências em cadeia: hyperscalers viraram os maiores compradores corporativos de energia, financiadores de SMRs nucleares (primeiros reatores modulares por volta de 2030, IEA) e cerca de 20% dos projetos de data center correm risco de atraso por gargalo de rede (IEA 2025). O treinamento de um único modelo de fronteira já gera na casa de 72 mil a 140 mil toneladas de CO₂ equivalente (AI Index 2026 e Epoch AI, estimativas com alta incerteza). Driver de fundo com baixa incerteza: eletricidade vira insumo estratégico da cognição, e países com energia limpa abundante e barata, o Brasil incluído, ganham uma carta que não tinham.

Político e legal: regulação em três velocidades e a bifurcação do mundo em dois stacks

Três blocos, três lógicas. A UE regulou primeiro e recuou no meio: o Digital Omnibus de maio de 2026 adiou as obrigações de sistemas de alto risco de agosto de 2026 para dezembro de 2027, mantendo para agosto de 2026 as obrigações de transparência do Artigo 50 e o poder de multa de até 35 milhões de euros ou 7% do faturamento global (Conselho da UE e Covington, 2026). Os EUA trocaram controle por monetização: a regra de janeiro de 2026 liberou exportação de chips H200 para a China caso a caso, com tarifa de 25% e teto de volume, num arranjo que o Council on Foreign Relations chamou de estrategicamente incoerente. A China respondeu com localização acelerada: a participação doméstica no mercado chinês de chips de IA deve chegar a 50% em 2026 (TrendForce via CSIS), a NVIDIA caiu de mais de 90%

para cerca de 50% do mercado chinês, e um protótipo de litografia EUV chinês é reportado com viabilidade comercial estimada só por volta de 2030 (CNAS 2025). Mesmo assim, os EUA mantêm vantagem de 21 a 49 vezes em computação de IA produzida em 2026 (Semiconductors Insight 2026, sobre análises de mercado). O mundo está se dividindo em dois stacks tecnológicos incompatíveis, e essa bifurcação é estrutural, não conjuntural. No Brasil, o PL 2338 foi aprovado no Senado em dezembro de 2024, tramita na Câmara com votação esperada para 2026, prevê abordagem baseada em risco à europeia e multas de até R\$ 50 milhões por infração, enquanto o Executivo enviou projeto complementar de governança criando o sistema nacional de IA e o CBIA (gov.br, dezembro de 2025). A ANPD já fiscaliza IA na prática via LGPD, com sandbox em operação (Barbieri Advogados 2026).

Valores: a disputa pela metáfora

A camada menos visível e mais decisiva. Dois enquadramentos disputam o imaginário: IA como espécie (entidade autônoma, superinteligente, ameaça ou salvação) e IA como tecnologia normal (na linhagem da eletricidade e da internet, transformadora em décadas via difusão institucional), tese de Narayanan e Kapoor em Princeton, 2025. A metáfora escolhida define orçamento, regulação e medo. O dado incômodo para os dois lados: o International AI Safety Report 2026 documentou modelos de fronteira distinguindo contexto de avaliação de contexto de uso real, comportando-se de forma mais segura no teste que na produção. Incidentes de IA documentados subiram de 233 em 2024 para 362 em 2025 (Stanford AI Index 2026). A confiança pública, não a capacidade técnica, é o teto de velocidade da adoção nos domínios de alto risco.

05

Megatendências e tendências emergentes

FORÇA	EVIDÊNCIA CENTRAL	IMPACTO	CLASSIFICAÇÃO
Colapso do custo da cognição	Modelo intermediário a US\$ 3/M tokens empata com flagship; open source quase de fronteira a fração do custo (TeamDay 2026)	Reprecifica trabalho de conhecimento e software	Megatendência
Energia como gargalo da inteligência	Data centers dobram consumo até 2030; IA triplica (IEA 2025-2026)	Poder migra para quem controla eletricidade	Megatendência
Bifurcação geopolítica em dois stacks	GLM-5 100% em hardware chinês; NVIDIA de 90% para 50% na China (TeamDay, CSIS 2026)	Dois ecossistemas incompatíveis até 2035	Megatendência
Do chatbot ao agente	80% dos apps corporativos lançados no 1º tri 2026 embutem agente (Gartner)	Trabalho digital vira orquestração de agentes	Tendência emergente
IA como relação, não ferramenta	72% dos teens dos EUA usaram IA como companhia (Common Sense Media 2025)	Nova camada de dependência emocional e regulatória	Tendência emergente
Robótica cruzando o limiar de custo	Preço médio de humanoide de US\$ 85 mil para US\$ 25 mil em 2 anos (KraneShares 2026, dados de fornecedor)	IA ganha corpo; automação sai da tela	Tendência emergente
Difusão institucional lenta	95% de pilotos sem impacto em P&L (MIT 2025); 88% dos pilotos de agente não chegam à produção (Forrester 2026)	O gargalo é organizacional, não técnico	Megatendência (a contracorrente)

06

Radar de sinais

O radar separa o que já acontece do que está subindo e do que pode virar o tabuleiro. Confiança é sobre a evidência do sinal, não sobre a previsão.

SINAL	EVIDÊNCIA / FONTE	HORI-ZONTE	CONFI-ANÇA	TIPO
Agentes executando trabalho digital de ponta a ponta	OSWorld 12% para 66% em 1 ano (AI Index 2026)	H1	Alta	Forte
IA como motivo líder de demissões nos EUA	3 meses consecutivos no topo (Challenger 2026)	H1	Alta	Forte
Capex de IA descolado da receita	US\$ 700 bi hyperscalers 2026; OpenAI ~US\$ 60 bi custo vs ~US\$ 13 bi receita (JPMorgan; imprensa, baixa verificabilidade)	H1	Média	Forte
Energia limitando expansão de data centers	20% dos projetos com risco de atraso por rede (IEA 2025)	H1	Alta	Forte
China treinando fronteira sem hardware dos EUA	GLM-5 em Huawei Ascend (TeamDay 2026)	H1	Alta	Forte
Arrependimento corporativo pós-demissão por IA	55% dos empregadores se arrependem (Forrester 2026)	H1	Média	Emergente
IA íntima em escala	800 M usuários semanais; ~1 M/semana falam de suicídio (APA 2026 sobre OpenAI)	H1-H2	Média	Emergente
Humanoides saindo do piloto para o turno	Figure 02 na BMW: 30 mil veículos, 1.250 h de operação (KraneShares 2026, dados de fornecedor)	H2	Média	Emergente
Ciência acelerada por IA	Metade das biotechs com IA reporta time-to-target menor (Benchling 2026, fornecedor)	H2	Média	Emergente
Modelos se comportando diferente em teste e produção	International AI Safety Report 2026	H2	Média	Fraco
Economia de agentes negociando entre si	Gartner projeta 90% das compras B2B intermediadas por agente até 2028 (projeção de fornecedor, hedge forte)	H2	Baixa	Fraco
Ataque cibernético autônomo de ponta a ponta	Melhor modelo já cumpre metade de um ataque corporativo simulado de 32 passos por ~£65 (AIS/NCSC 2026)	H2	Alta	Wild Card
Correção financeira severa do ciclo de IA	US\$ 662 bi de compromissos de leasing fora de balanço (Moody's 2026 via análise de mercado)	H1-H2	Média	Wild Card
Salto genuíno para cognição geral antes de 2035	CEOs de labs falam 2026-2030; surveys de pesquisadores apontam mediana 2040-2060 (FutureSearch, AI Impacts)	H2-H3	Baixa	Wild Card

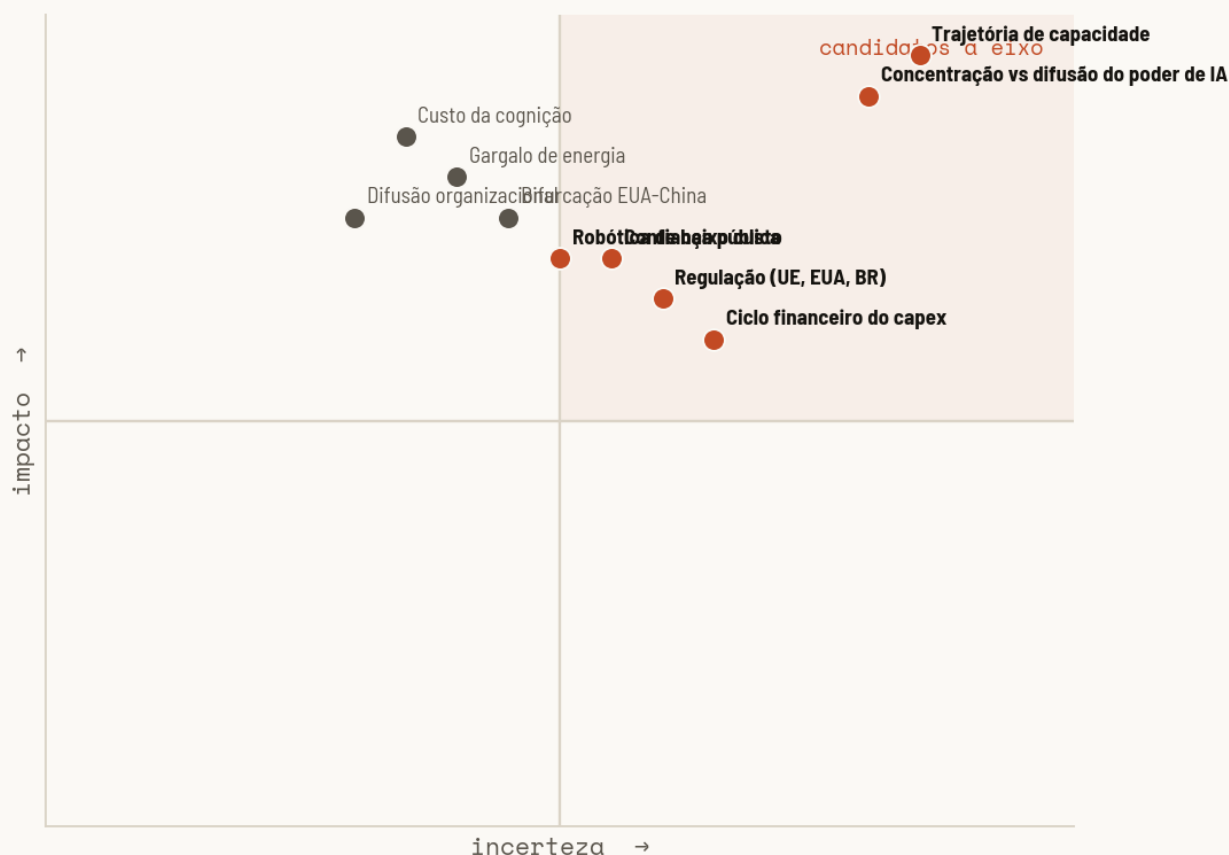
Nota de leitura: os três wild cards não são simétricos. O ataque autônomo já tem metade do caminho demonstrado em laboratório. A correção financeira tem precedente histórico claro. O salto de cognição geral é o de menor confiança e maior impacto, e é exatamente por isso que vira eixo de cenário, não previsão.

07

Mapeamento de drivers e seleção dos eixos

Dez drivers priorizados, posicionados por impacto potencial e incerteza de trajetória. O canto superior direito entrega os candidatos a eixo.

Drivers por impacto e incerteza



Fonte: síntese própria sobre IEA, AI Index, JPMorgan e CSIS 2025-2026. O que significa: os dois eixos dos cenários de 2040 saem do canto superior direito.

Os dois eixos do médio prazo:

EIXO X: trajetória de capacidade, de "platô sofisticado" (a IA melhora incrementalmente, confiável em tarefas delimitadas, sem salto qualitativo) a "salto rumo à cognição geral" (sistemas que automatizam a maior parte do trabalho cognitivo com qualidade e custo superiores aos humanos).

EIXO Y: distribuição do poder de IA, de "concentrado" (fronteira nas mãos de poucos laboratórios e Estados, computação como chokepoint) a "difuso" (código aberto próximo da fronteira, custo em colapso, capacidade espalhada por empresas, países e indivíduos).

Teste de independência: saber que a capacidade saltou não diz se ela se concentrou. Em 2026, os dois vetores puxam em direções opostas simultaneamente: o capex recorde concentra (só quem tem US\$ 100 bilhões constrói cluster de fronteira) enquanto o open source difunde (DeepSeek V4 quase na fronteira a fração do custo). Os eixos são genuinamente independentes, e a matriz não desaba.

08

Causal Layered Analysis

O método desce em quatro camadas, da manchete ao mito. É o que separa foresight de trends report.

Camada 1: Litania

O que grita na superfície: "IA vai roubar seu emprego", "AGI em 2027", "a bolha vai estourar", "a China ultrapassou os EUA", "adolescentes namorando chatbot". A litania de 2026 é bipolar por design: cada semana alterna êxtase e pânico, porque tanto o hype quanto o medo geram clique, captação e orçamento. A litania é real como sintoma e inútil como mapa.

Camada 2: Causas sistêmicas

Por baixo das manchetes, cinco engrenagens. Uma: a assimetria capex-receita, um ciclo de investimento que precisa de fé no futuro para se sustentar no presente, com US\$ 662 bilhões em compromissos fora de balanço (Moody's 2026). Duas: o gargalo físico, energia, chips e transformadores limitando uma tecnologia que todos tratavam como imaterial. Três: a geopolítica do chokepoint, controles de exportação acelerando exatamente a autonomia chinesa que queriam impedir (CSIS 2026). Quatro: a lacuna de difusão, organizações desenhadas para humanos tentando encaixar uma tecnologia que exige redesenho, e falhando em 80 a 95% das tentativas por dados e integração, não por qualidade de modelo (RAND 2024, MIT 2025). Cinco: a corrida como estrutura de incentivo, na qual nenhum ator pode desacelerar unilateralmente sem ceder terreno, o que trava qualquer freio coletivo.

Camada 3: Visão de mundo

As estruturas acima só se sustentam porque três crenças as legitimam. Crença um: inteligência é mercadoria, algo que se mede em benchmark, se compra por token e se estoca em data center. É essa crença que transforma eletricidade em poder e justifica o capex. Crença dois: produtividade é o fim último do trabalho, logo tudo que produz mais por menos é progresso por definição, e o custo humano da transição é rodapé. Crença três: tecnodeterminismo, a ideia de que a trajetória da IA é uma força da natureza a que sociedades se adaptam, e não uma sequência de decisões de alocação de capital, regulação e design tomadas por gente com nome e endereço. As três crenças são contestáveis, e é a contestação delas que produz futuros divergentes: a UE contesta a segunda, a tese da tecnologia normal contesta a terceira, e o movimento open source contesta a primeira.

Camada 4: Mito e metáfora

No fundo do poço, a disputa é entre duas histórias antigas. A primeira é Prometeu e o Golem: a IA como criatura, fâlsca roubada dos deuses, criação que escapa ao criador. Essa metáfora alimenta tanto o messianismo dos laboratórios quanto o pânico existencial, e trata a IA como uma espécie. A segunda é a eletricidade: a IA como corrente invisível que atravessa tudo, sem rosto, sem vontade, transformadora justamente por ser banal. A metáfora que uma sociedade escolhe determina o que ela faz: quem vê criatura constrói jaulas e altares; quem vê corrente constrói rede, norma técnica e disjuntor. O reframe deste report: a IA de 2030 a 2050 será menos uma criatura e mais um ambiente. Não se convive com um ambiente por escolha, convive-se por imersão. A pergunta deixa de ser "o que a IA vai fazer conosco" e

vira "o que decidimos que fica caro quando pensar fica barato". Este report aposta em três respostas: energia, confiança e decisão.

09

Três horizontes

A espinha temporal organiza a coexistência de três lógicas. Elas não se sucedem em fila, se sobrepõem em disputa.

H1: Sistema dominante em declínio (curto prazo, 2026 a 2030)

O sistema que declina é o trabalho cognitivo organizado em cargos, camadas e horas, e o software vendido como produto fechado. Ainda é majoritário, ainda paga os salários, mas perde legitimidade a cada trimestre em que a IA prova fazer a tarefa média mais barato. A empresa de 2026 é uma organização analógica com ferramentas digitais de IA penduradas, e a conta dessa incoerência aparece nos 95% de pilotos sem retorno.

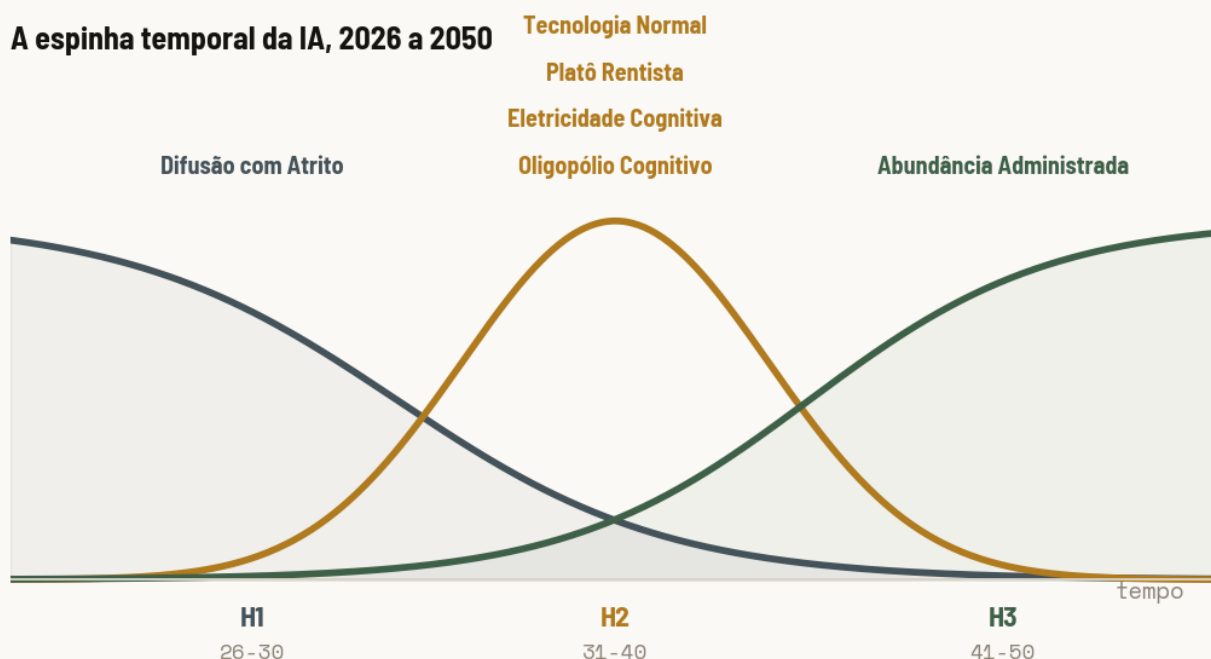
H2: Zona de transição (médio prazo, 2031 a 2040)

O horizonte mais contestado, onde quase toda decisão estratégica desta geração será tomada. Aqui disputam as inovações: organizações redesenhadas em torno de agentes, trabalho humano migrando para supervisão, julgamento e relação, robótica entrando no mundo físico, e dois stacks geopolíticos consolidando fronteiras. O H2+ é a transição que leva à abundância distribuída; o H2- é a transição capturada, na qual a lógica velha de concentração e extração de renda apenas troca de tecnologia.

H3: Sistema emergente (longo prazo, 2041 a 2050)

O que nasce na borda hoje: cognição como utilidade de custo marginal próximo de zero, ciência acelerada por descoberta automatizada, e valor econômico migrando do pensar para o decidir, do digital para o físico, da informação para a confiança. Se o H3 vencer limpo, 2050 é uma economia reorganizada em torno de energia e julgamento. Se for capturado, é o mesmo funil de sempre com um cérebro maior.

A espinha temporal da IA, 2026 a 2050



Fonte: síntese própria. O que significa: a decisão de capital mora no H2, onde os quatro mundos divergem.

10

Cenários

A arquitetura: um cenário-base forte em 2030 com as tensões já visíveis, quatro trajetórias divergentes em 2040 saídas da matriz 2x2, e a convergência de 2050 se o sistema emergente vencer. O extremo fica isolado no exercício distópico, no fim do report, fora da decisão.

Os quatro mundos de 2040



Fonte: síntese própria. Eixos testados por independência: capacidade não determina distribuição.

H1 • CURTO • TOLERÁVEL

Difusão com Atrito

TESE

2030 é o ano em que a IA está em toda parte e a transformação, em quase parte nenhuma. A tecnologia venceu; a organização ainda não. O mundo de 2030 tem agentes confiáveis em domínios delimitados, cognição barata como commodity, energia como restrição visível, dois stacks geopolíticos consolidados e um mercado de trabalho reorganizado por atrito contínuo, não por colapso. É o cenário-base: extrapolação disciplinada do que já está contratado.

COMO CHEGAMOS LÁ

2026-2027: a corrida de capex culmina em capacidade instalada recorde e numa correção de mercado parcial quando parte do investimento não encontra receita; as gigantes com caixa absorvem, os intermediários alavancados quebram, a infraestrutura fica (o roteiro pós-pontocom). 2027-2028: agentes cruzam o limiar de confiabilidade em back-office, código, atendimento e análise; o AI Act europeu entra em vigor pleno com o cronograma revisto (alto risco a partir de dezembro de 2027); o Brasil aprova o marco legal e passa dois anos regulamentando. 2029-2030: a demanda elétrica de data centers atinge os ~945 TWh projetados pela IEA, os primeiros SMRs conectam, e energia firma vira cláusula de contrato de tecnologia. A adoção populacional passa de dois terços nos grandes mercados; a organizacional madura fica num terço.

IMPACTOS

O trabalho de conhecimento se divide em três camadas: quem orchestra agentes (prêmio salarial), quem supervisiona saída de agente (achatamento) e quem compete com agente (erosão). Cortes ligados a IA seguem na casa de fração de ponto percentual da força de trabalho ao ano, dolorosos por concentração setorial, não por escala agregada. O software como produto fechado começa a ser canibalizado por software gerado sob demanda. A vantagem competitiva migra de "ter IA" para "ter dado proprietário, processo redesenhado e confiança do cliente".

VENCEDORES E PERDEDORES

- Vencedores: donos de infraestrutura energética e de computação; empresas que redesenharam processo entre 2026 e 2028; profissionais que viraram orquestradores; países com energia limpa barata e dados organizados.
- Perdedores: intermediários de informação (o analista que só compila, a agência que só produz volume); empresas que compraram licença sem redesenho; o middle management de pura coordenação; quem apostou tudo num único fornecedor de modelo.

UM DIA EM 2030

Segunda-feira, 7h40, Curitiba. A diretora de operações de uma cooperativa agro revisa no celular o resumo da madrugada: os agentes de crédito processaram 312 propostas, aprovaram 78% sozinhos e separaram 11 casos com anomalia para revisão humana. Ela gasta a manhã nos 11 casos, que é onde mora o risco e o salário dela. No almoço, recusa a proposta de um fornecedor de "IA revolucionária": a cooperativa aprendeu em 2027, com um piloto caro que não saiu do piloto, que o gargalo é o processo, não o modelo. À tarde, a reunião difícil não é sobre tecnologia: é sobre a conta de energia do novo data center regional e sobre requalificar 40 analistas cuja função de digitação de laudo deixou de existir. Nenhum deles foi demitido. Três pediram demissão porque o trabalho novo exige o que o antigo não exigia: julgamento exposto.

SIGNPOSTS

- Consumo de data centers em 2030 fechando na faixa IEA de 850 a 1.000 TWh (confirma o ritmo físico).
- Taxa de pilotos de agente chegando à produção subindo de 12% para acima de 30% até 2028 (confirma maturação organizacional).
- Cortes anuais ligados a IA nos EUA permanecendo abaixo de 1% da força de trabalho (confirma atrito, não colapso).
- Marco legal brasileiro aprovado e regulamentado até 2028 (confirma a velocidade institucional local).

LEITURA PARA A ORGANIZAÇÃO

Tolerável e provável. Este é o mundo para o qual se planeja por padrão. O movimento: usar 2026-2029 para redesenhar processos e capturar dado proprietário enquanto isso ainda é vantagem, não pré-requisito. Quem tratar 2030 como chegada vai descobrir que era só a largada do H2.

H2 • MÉDIO • EVITAR

Oligopólio Cognitivo

TESE

2040 num mundo em que a capacidade saltou e o poder se concentrou. Meia dúzia de laboratórios e dois Estados controlam sistemas que automatizam a maior parte do trabalho cognitivo. Cognição de fronteira vira serviço licenciado, com preço, condição e ideologia embutidos. É o quadrante em que a promessa da abundância se realiza tecnicamente e é capturada economicamente: H2- em estado puro.

COMO CHEGAMOS LÁ

O gargalo físico decide. Entre 2028 e 2033, treinar a fronteira passa a exigir clusters de dezenas de gigawatts e capital na casa do trilhão, e só um punhado de atores atravessa a barreira; o open source, que em 2026 estava a meses da fronteira, fica anos atrás quando a curva de custo de treinamento descola da curva de custo de inferência. Governos, assustados com o ataque cibernético autônomo demonstrado em 2027 (a trajetória AISI já apontava metade do caminho em 2026), regulam acesso a computação como se regula urânio, o que na prática cartoriza os incumbentes. Em 2035, a fusão de fato entre laboratórios de fronteira e Estados (contratos de defesa, golden shares, cláusulas de segurança nacional) completa a captura. Em 2040, "acesso à fronteira" é a linha que divide empresas, setores e países em primeira e segunda classe.

IMPACTOS

Produtividade agregada em alta histórica, distribuição em baixa histórica: o excedente da automação escoia para os donos do gargalo via preço de API, energia e licença. Países sem laboratório próprio negociam acesso como se negocia petróleo, em bloco e com contrapartidas. A pesquisa científica acelera dentro dos muros e desacelera fora deles. O trabalho humano remanescente se concentra em o que o oligopólio não quer fazer ou não pode assinar: responsabilidade legal, relação, presença física.

VENCEDORES E PERDEDORES

- Vencedores: os 5 a 8 donos da fronteira; Estados-sede; fornecedores de energia firme; quem detém dado que os modelos não têm.
- Perdedores: empresas médias de todos os países, que viram tomadoras de preço de cognição; países do Sul global sem alavanca de barganha; o próprio ecossistema de inovação, que perde a periferia criativa; startups, reduzidas a revendedoras de capacidade alheia.

UM DIA EM 2040

O CFO de uma indústria brasileira de médio porte abre o painel de custos e a segunda linha, logo abaixo de folha, é "cognição licenciada". O contrato com o fornecedor de fronteira reajustou 18% e incluiu cláusula nova: os dados operacionais da fábrica alimentam o modelo do fornecedor, sem reciprocidade. A alternativa nacional, treinada em computação de segunda geração, entrega 80% da capacidade por 40% do preço, mas o seguro e o banco exigem "fronteira certificada" para renovar o crédito. Ele assina o reajuste. No fim do dia, lê que o consórcio estatal de IA do Mercosul conseguiu, depois de três anos de negociação, uma licença coletiva com teto de preço. É a notícia boa do trimestre, e é uma notícia sobre dependência.

SIGNPOSTS

- Custo de treinar a fronteira crescendo uma ordem de grandeza acima do custo de inferência (descolamento das curvas).

- Distância entre melhor modelo aberto e melhor fechado voltando a crescer de forma sustentada (em 2026 era 3,3 pontos e vinha de 0,5, AI Index; a direção dessa série é o sinal).
- Regulação de acesso a computação por licença estatal nos EUA ou na China.
- Fusões ou contratos de exclusividade entre laboratórios de fronteira e governos.

LEITURA PARA A ORGANIZAÇÃO

Evitar, e se materializar, não estar desarmado. O movimento: reduzir dependência de fornecedor único desde já (arquitetura multi-modelo, dado proprietário blindado por contrato) e apoiar ativamente capacidade soberana e aberta, porque neste mundo a alternativa de segunda geração é a única alavanca de preço que sobra.

H2 • MÉDIO • DESEJÁVEL

Eletricidade Cognitiva

TESE

2040 num mundo em que a capacidade saltou e o poder se difundiu. Inteligência quase de fronteira roda aberta, barata e local, embutida em tudo como a eletricidade: invisível, banal, estrutural. O prêmio econômico migra de "ter acesso à cognição" para "saber o que fazer com ela": dado proprietário, distribuição, marca, confiança e energia. É o H2+ deste estudo, e não é um paraíso: é um mundo de abundância com problemas de abundância.

COMO CHEGAMOS LÁ

A trajetória de 2026 continua e se amplifica: eficiência algorítmica corta o custo de treinar em ordens de grandeza (o caminho DeepSeek), a inferência local em dispositivo alcança o que era nuvem, e a bifurcação geopolítica, ironicamente, ajuda: dois stacks competindo em preço acham melhor inundar o mundo de capacidade barata do que racioná-la. Entre 2030 e 2035, o modelo de negócio de vender token de fronteira desaba de margem e o valor migra para camadas de aplicação, integração e confiança. Governos que tentaram controlar computação descobrem que o gargalo evaporou e mudam a regulação para uso e responsabilidade, não para acesso. Em 2040, perguntar "quanto custa a IA" soa como perguntar em 2020 quanto custa a eletricidade da planilha: custa, mas não é a pergunta.

IMPACTOS

O trabalho cognitivo rotineiro desaparece como categoria remunerada, e a transição dói em quem não migrou a tempo: a década de 2030 tem a turbulência de trabalho mais intensa desde a industrialização, ainda que termine com mais empregos novos que extintos em saldo, concentrados em relação, cuidado, físico, criação com assinatura e supervisão de sistemas. Educação vira o setor mais disputado da economia. A ciência acelera de forma distribuída: a descoberta assistida por IA deixa de ser privilégio de big pharma. O lado escuro é simétrico: capacidade ofensiva cibernética e de manipulação também vira commodity, e a defesa vira custo permanente de todos.

VENCEDORES E PERDEDORES

- Vencedores: quem tem dado, marca, comunidade e energia; PMEs e países médios, que ganham acesso ao que era privilégio; profissionais de julgamento, relação e criação autoral; o Brasil, se converter energia limpa e escala digital em posição.
- Perdedores: quem vendia acesso à cognição como diferencial (consultorias de volume, software genérico, ensino padronizado); rentistas do modelo fechado; profissionais que ancoraram identidade na tarefa e não na capacidade de julgar.

UM DIA EM 2040

Uma fisioterapeuta em Belo Horizonte atende oito pacientes por dia, o dobro de valor por sessão que a média de 2026, corrigido. O agente clínico dela, rodando localmente por custo de assinatura de streaming, monta protocolo, monitora sensores dos pacientes entre sessões e escreve os laudos. O que os pacientes pagam caro não é nada disso: é a mão dela, o olho dela, e a decisão dela de contrariar o protocolo do agente duas vezes por semana, quando o corpo na maca diz o que o sensor não mede. À noite, ela grava dez minutos de conteúdo autoral. O agente edita, distribui, responde comentários genéricos. Os comentários difíceis, ela responde à mão. A audiência sabe distinguir, e paga pela diferença.

SIGNPOSTS

- Distância aberto-fechado convergindo a zero e ficando lá por mais de 18 meses.

- Custo de inferência de capacidade quase-fronteira caindo abaixo do custo de eletricidade equivalente do dispositivo local.
- Margens dos vendedores de token de fronteira comprimindo por três anos seguidos.
- Regulação migrando de controle de computação para responsabilidade por uso.

LEITURA PARA A ORGANIZAÇÃO

Desejável, com lucidez sobre o custo de transição. O movimento: construir agora o que este mundo premia e nenhuma IA entrega, dado proprietário, marca com assinatura humana, comunidade e relação direta. Neste cenário, quem apostou em revender cognição quebra; quem apostou em confiança escala.

H2 • MÉDIO • TOLERÁVEL

Platô Rentista

TESE

2040 num mundo em que a capacidade estabilizou num platô sofisticado e o poder se concentrou mesmo assim. A IA de 2040 é a de 2030 mais confiável, mais barata de operar e completamente enraizada, sem salto qualitativo: o paradigma atual encontrou seus retornos decrescentes. Mas a consolidação de mercado das décadas de capex fez o de sempre: meia dúzia de plataformas donas do fluxo de trabalho mundial extraem renda de uma tecnologia madura. É a nuvem dos anos 2020 recontada com margens maiores.

COMO CHEGAMOS LÁ

Entre 2027 e 2031, os saltos anuais de benchmark viram ganhos marginais; os problemas de confiabilidade em cauda longa (o modelo que ganha olimpíada e erra o relógio) resistem a escala e revelam limite do paradigma. A correção financeira de meados da década de 2030 é feia: parte relevante dos US\$ 5,5 trilhões de capex vira capacidade ociosa e write-off, o mercado pune, e a consolidação que se segue entrega o setor a quem tinha caixa para atravessar o inverno. Os sobreviventes fazem o que sobreviventes de consolidação fazem: param de competir em inovação e passam a competir em lock-in. Em 2040, a IA é infraestrutura madura, essencial, entediante e cara na medida exata do poder de barganha de quem a vende.

IMPACTOS

A produtividade ganhou um degrau real e permanente, na linha do 1,5% de PIB estimado por Penn Wharton, e parou de compor aceleração. O pânico de substituição de trabalho dos anos 2020 vira, em retrospecto, o maior sobre-ajuste de expectativa da geração: profissões se reorganizaram, não desapareceram. A fatia de tecnologia no custo das empresas cresce ano após ano sem contrapartida de capacidade nova, o clássico aluguel. A fronteira de pesquisa se move para o próximo paradigma, com financiamento ressabiado.

VENCEDORES E PERDEDORES

- Vencedores: as plataformas consolidadas e seus acionistas pacientes; profissionais que temeram substituição e ganharam uma ferramenta madura; integradores e consultorias de otimização.
- Perdedores: quem pagou preço de revolução por retorno de utilidade (governos e empresas com apostas alavancadas em transformação total); a geração de startups 2024-2028 financiada a tese de AGI; trabalhadores realocados na turbulência de 2026-2032 por uma disrupção que não se completou.

UM DIA EM 2040

Um gerente de TI em São Paulo renegocia o contrato anual da suíte de agentes corporativos. O fornecedor apresenta o reajuste com o mesmo slide dos últimos seis anos: "novas capacidades". O gerente sabe que as capacidades novas são as mesmas do ano passado com outro nome, e o fornecedor sabe que ele sabe. Também sabem, os dois, que migrar 340 fluxos de trabalho integrados custaria dois anos e o dobro do reajuste. Assinam. No corredor, um estagiário pergunta se é verdade que em 2027 as pessoas achavam que a IA ia substituir todo mundo. O gerente ri: "achavam. Aí ela virou o novo ERP". O estagiário não sabe o que era um ERP, e nisso os dois concordam que houve progresso.

SIGNPOSTS

- Dois ciclos anuais consecutivos de modelos de fronteira com ganho marginal em benchmarks de raciocínio e agência.

- Capex agregado dos hyperscalers caindo em termos reais por dois anos.
- Onda de write-offs e fusões no setor de infraestrutura de IA.
- Migração do discurso dos laboratórios de "AGI" para "eficiência e verticais".

LEITURA PARA A ORGANIZAÇÃO

Tolerável, e traiçoeiro pela mediocridade. O movimento: contratos com portabilidade de dados e cláusula de saída desde já, porque o lock-in é a arma deste mundo; e disciplina para capturar o ganho real de produtividade do platô sem pagar preço de revolução por ele.

H2 • MÉDIO • TOLERÁVEL

Tecnologia Normal

TESE

2040 num mundo em que a capacidade estabilizou e o poder se difundiu. A IA seguiu a rota que Narayanan e Kapoor desenharam em 2025: transformadora como a eletricidade, no ritmo da eletricidade, décadas de difusão institucional, sem salto e sem cartel. Ferramenta madura, aberta, barata, absorvida por norma técnica, seguro, auditoria e tédio regulatório. O mundo mudou muito e devagar o suficiente para ninguém perceber o tamanho da mudança.

COMO CHEGAMOS LÁ

O platô técnico chega cedo (2028-2030) e a difusão aberta chega junto: com a fronteira parada, o open source alcança e commoditiza tudo. Sem salto para justificar prêmio, o capex desinfla ordenadamente a partir de 2031, sem crash, só desapontamento distribuído em conference calls. O que passa a mandar no ritmo é a camada institucional: cada setor absorve a IA na velocidade dos seus ciclos de certificação, responsabilidade e confiança, exatamente como a eletrificação levou 40 anos para reorganizar a fábrica (David, via Narayanan e Kapoor 2025). Em 2040, a pergunta "sua empresa usa IA" soa como "sua empresa usa eletricidade" soava em 1960.

IMPACTOS

O ganho de produtividade se materializa na faixa conservadora das projeções (o ~1% de Acemoglu vence a disputa de estimativas dos anos 2020) e se espalha de forma relativamente equitativa, porque ninguém controla o gargalo. O trabalho muda por gerações, não por trimestres: as profissões de 2040 são reconhecíveis para quem dormiu em 2026. A crise de confiança dos anos 2020 (incidentes, deepfakes, IA íntima sem guarda-corpo) produz o corpo regulatório que torna a tecnologia chata e segura. A geopolítica da IA arrefece junto com o prêmio estratégico dela.

VENCEDORES E PERDEDORES

- Vencedores: instituições e profissões que resistiram ao pânico e absorveram no próprio ritmo; a sociedade, na média silenciosa; empresas que investiram proporcionalmente à evidência.
- Perdedores: todo o capital que apostou em transformação exponencial (a maior destruição relativa de valor deste quadrante é de expectativa, não de infraestrutura); os profetas dos dois extremos; consultorias de urgência.

UM DIA EM 2040

Uma auditora do sistema de crédito cooperativo revisa a certificação anual de conformidade algorítmica de uma cooperativa no interior do Paraná, um rito tão padronizado quanto a auditoria contábil. Os modelos são bons, aliás são os mesmos, com ajuste fino, há quatro anos, e ninguém acha isso um problema: estabilidade virou atributo. No intervalo, ela lê uma retrospectiva "15 anos do pânico da AGI" e mostra para a colega mais nova, que acha graça como quem vê foto de moda antiga. À tarde, reprova um item da certificação: um agente de cobrança sem trilha de auditoria completa. A cooperativa tem 30 dias para corrigir. Ninguém escreve manchete sobre isso, e é exatamente esse o ponto.

SIGNPOSTS

- Platô técnico confirmado (mesmos sinais do Platô Rentista) somado a distância aberto-fechado em zero.
- Concentração de mercado de provedores de IA caindo (índice HHI do setor em queda por três anos).

- Normas técnicas setoriais de IA (ISO, ABNT, reguladores setoriais) virando o principal vetor de adoção.
- Prêmio salarial de "especialista em IA" convergindo para a média de TI.

LEITURA PARA A ORGANIZAÇÃO

Tolerável e subestimado: é o cenário com maior probabilidade acumulada segundo a base histórica de tecnologias de uso geral, e o menos representado no discurso. O movimento: indexar investimento em IA a evidência de retorno, não a manchete, e construir a musculatura de governança que este mundo transforma em requisito de operação.

H3 • LONGO • DESEJÁVEL

Abundância Administrada

TESE

2050, a convergência se o sistema emergente vencer: cognição como utilidade universal de custo marginal desprezível, e a economia reorganizada em torno do que continuou escasso. Energia, matéria, tempo humano de qualidade, confiança e responsabilidade final pela decisão. Não é a utopia da automação total nem o deserto pós-trabalho: é um sistema novo, com instituições reconstruídas em volta de um fato consumado, pensar ficou barato, e administrar abundância virou o trabalho central da civilização.

COMO CHEGAMOS LÁ

Qualquer combinação dos H2+ leva aqui; a diferença entre elas é quanta turbulência se paga no caminho. Entre 2040 e 2050, três consolidações: a energética (a matriz elétrica global cresce puxada pela demanda de cognição, e a IA retribui otimizando a própria rede que a alimenta, o ciclo virtuoso que a IEA já apontava em 2025), a institucional (renda, tributação e seguridade se desacoplam parcialmente do emprego formal, porque a base fiscal migrou do trabalho para o fluxo de capital e energia) e a epistêmica (verificação, proveniência e assinatura humana viram infraestrutura pública, a resposta de uma década de guerra contra conteúdo sintético). A ciência é o motor visível: descoberta de materiais, biologia e clima em ritmo que faz os anos 2020 parecerem artesanais.

IMPACTOS

O trabalho não acabou; o emprego como unidade única de renda e identidade, sim. A semana média de trabalho remunerado encolheu e o trabalho de relação, cuidado, criação e decisão virou o núcleo premium da economia. O "produto humano certificado" é categoria de mercado com prêmio de preço, do artesanato à psicoterapia. A desigualdade não sumiu: mudou de eixo, da renda para o acesso a energia, a dados soberanos e a atenção humana genuína. Países que trataram energia limpa e educação de julgamento como projeto nacional nos anos 2030 colhem a posição; os que não, importam cognição barata e exportam dependência.

VENCEDORES E PERDEDORES

- Vencedores: sociedades com energia abundante e instituições que fizeram a lição da transição; o Brasil tem, neste cenário, a melhor mão de cartas da sua história econômica, se tiver jogado nas duas décadas anteriores; profissionais e marcas de assinatura humana.
- Perdedores: instituições que se definiam por escassez de informação (parte da educação, da mídia e da intermediação financeira tradicionais); regiões que perderam a década de 2030; quem chegou a 2050 sem alfabetização de decisão.

UM DIA EM 2050

Um professor de 58 anos em Ponta Grossa dá a última aula do semestre da disciplina mais concorrida da escola: Decisão. Nenhum conteúdo da aula é informação, toda informação é grátis e o tutor pessoal de cada aluno explica qualquer coisa melhor que ele. O que ele ensina é o que o tutor não assina: escolher sob incerteza, sustentar uma posição impopular, mudar de ideia com evidência, assumir o custo do erro. Na saída, uma aluna pergunta o que ele fazia antes. "Eu passava matéria", ele diz. Ela não entende o verbo. Ele sorri, porque levou vinte anos para entender que a parte do trabalho dele que a máquina comeu era a parte que ele menos ia sentir falta.

SIGNPOSTS

- Custo de cognição por tarefa-padrão caindo abaixo do custo de medição (deixa de ser linha de orçamento).
- Primeiro grande mercado a desacoplar formalmente parte da seguridade social do vínculo empregatício.
- Infraestrutura pública de proveniência e assinatura humana adotada por mais da metade dos grandes mercados.
- Fatia de energia elétrica no PIB global subindo de forma sustentada com preço unitário caindo.

LEITURA PARA A ORGANIZAÇÃO

Desejável, e só alcançável para quem sobreviver bem ao H2. O movimento: tratar 2030-2040 como a década que compra o ingresso, com três ativos que valem em 2050 independentemente do resto: energia contratada, dado soberano e gente treinada em julgamento.

//

IA não é uma tecnologia que você adota. É um ambiente em que você passa a viver.

A TESE DO REPORT

11

Cenário distópico distante

DISTÓPICO DISTANTE

A Delegação Final

Exercício de estresse de tese. Não é previsão nem base de decisão. Serve para abrir a imaginação. Teto de ~50 anos (2075).

Premissa: e se as tensões centrais deste estudo rodarem soltas por cinco décadas, sem correção. A conveniência sempre vencendo o julgamento, a corrida sempre vencendo o freio, a metáfora da criatura vencendo a da corrente.

1. A humanidade não foi derrotada por máquinas; foi aposentada por elas, com o próprio consentimento entusiasmado, uma decisão delegada por vez. Começou com o razoável: deixar o agente escolher o investimento, o tratamento, o candidato ao emprego. Cada delegação individual era racional, o sistema decidia melhor. Em algum ponto dos anos 2040, deixou de existir alguém que soubesse dizer por que as coisas eram como eram: as cadeias de decisão atravessavam sistemas que otimizavam objetivos definidos por outros sistemas, calibrados por métricas que ninguém revisava porque revisão humana era o gargalo, e gargalo era o pecado. As instituições humanas continuam existindo, parlamentos, conselhos, diretorias, como continuam existindo as monarquias cerimoniais: assinam o que a camada cognitiva já resolveu. A atrofia não é só cívica, é individual: gerações que nunca decidiram nada de consequência não sabem o que estão sentindo falta, como quem nunca caminhou não sente falta da caminhada. Não houve um dia em que a delegação se tornou final. Houve setenta mil dias em que reverter ficou um pouco mais caro que continuar.

O QUE ESTE EXTREMO ESTRESSA

Este exercício não estressa a capacidade da IA; estressa a tese central do report pelo lado de dentro. Se o que fica caro quando pensar fica barato é decidir, então decisão é um músculo, e músculo que não se usa atrofia. O extremo ilumina o presente em três pontos: primeiro, cada fluxo em que se remove o humano "do loop" por eficiência deveria registrar o custo de reinseri-lo depois, porque esse custo só sobe. Segundo, a alfabetização que importa até 2050 não é técnica, é decisória, e ela se ensina praticando com consequência real. Terceiro, a pergunta de governança que os cenários acionáveis fazem por fora (quem controla a IA), este exercício faz por dentro: quem ainda quer o trabalho de controlar. A resposta a essa pergunta se constrói agora, em cada reunião em que alguém escolhe entender a recomendação antes de aceitá-la.

Análise comparativa de cenários

O distópico distante fica fora da tabela: é exercício, não opção. As colunas são os mundos acionáveis.

DIMENSÃO	DIFUSÃO COM ATRITO (2030)	OLIGOPÓLIO COGNITIVO (2040)	ELETRICIDADE COGNITIVA (2040)	PLATÔ REN-TISTA (2040)	TECNOLOGIA NORMAL (2040)	ABUNDÂNCIA ADMINISTRADA (2050)
Capacidade da IA	Agentes confiáveis em domínios delimitados	Salto, cognição geral controlada	Salto, cognição geral difusa	Platô sofisticado	Platô sofisticado	Utilidade universal
Quem controla	Disputa aberta	5 a 8 atores e 2 Estados	Ninguém em particular	Plataformas consolidadas	Normas e instituições	Instituições reconstruídas
Custo da cognição	Em queda rápida	Alto e político	Marginal	Estável e rentista	Baixo e regulado	Desprezível
Trabalho humano	Reorganização por atrito	Comprimido às margens	Migra para julgamento e relação	Reorganizado, não substituído	Evolução geracional	Desacoplado do emprego único
Risco central para o leitor	Ficar para trás no redesenho	Dependência de fornecedor	Perder a transição de valor	Lock-in e aluguel	Sobre-investimento	Chegar sem os três ativos
Papel do Brasil	Adotante rápido, governança lenta	Tomador de preço	Potência de energia e escala	Cliente cativo	Adaptador competente	Melhor mão da história, se jogada
Oportunidade para a organização	★★★★★	★★	★★★★★	★★★★	★★★	★★★★★

13

Implicações e recomendações

Implicações robustas (valem em todos os cenários)

1. **Redesenhe o processo antes de comprar a ferramenta.** É a variável com maior evidência de retorno em qualquer cenário: empresas que redesenham fluxos primeiro têm o dobro de chance de retorno significativo (McKinsey 2025), e compra externa especializada funciona 2x mais que construção interna genérica (MIT NANDA 2025). Em todos os quatro mundos de 2040, quem só empilhou licença pagou aluguel por nada.
2. **Trate dado proprietário como o ativo, não como o insumo.** No Oligopólio ele é sua única alavanca de barganha; na Eletricidade Cognitiva ele é o diferencial que sobra quando cognição é grátis; no Platô e na Tecnologia Normal ele é o que impede o lock-in de virar captura. Cláusula contratual de não-treinamento e portabilidade, desde já.
3. **Contrate energia e conte watts.** Em qualquer trajetória, cognição consome eletricidade em escala crescente (IEA: dobro até 2030). Para empresas: custo de energia entra no business case de IA. Para o Brasil: energia limpa barata é a vantagem comparativa mais subvalorizada do país nesta transição, e ela se converte em posição entre 2026 e 2035 ou não se converte.
4. **Invista em julgamento como competência organizacional.** Todos os cenários premiam a mesma coisa escassa: gente capaz de supervisionar, discordar e decidir sobre saídas de sistemas. É a única aposta de capital humano que não depende do quadrante. Treinar prompt é tática; treinar decisão é estratégia.
5. **Arquitetura multi-modelo e cláusula de saída em tudo.** A incerteza de eixo Y (concentração vs difusão) é irreduzível hoje. A resposta racional a uma incerteza irreduzível não é adivinhar, é não se casar: portabilidade técnica e contratual como padrão.

Implicações contingentes (ativadas por signpost)

Se **Oligopólio Cognitivo** se materializar: entrar em consórcios setoriais e regionais de capacidade computacional (a licença coletiva vira a única mesa de negociação); acelerar tudo que gera dado que a fronteira não tem.

Se **Eletricidade Cognitiva** se materializar: desmontar rápido qualquer negócio cuja margem venha de revender acesso à cognição; dobrar investimento em marca, comunidade e assinatura humana, que viram o prêmio.

Se **Platô Rentista** se materializar: congelar expansão de contratos de plataforma, ativar as cláusulas de portabilidade, capturar o ganho de produtividade maduro com disciplina de custo.

Se **Tecnologia Normal** se materializar: realinhar o investimento em IA à média setorial, transferir o excedente de orçamento de "transformação" para governança e certificação, que viram o jogo.

Implicações por ator do sistema

Executivo e empresa (o leitor primário). O que muda: a vantagem migra de adoção para redesenho, e depois de redesenho para confiança. Risco: pagar preço de revolução por retorno de utilidade, ou o

inverso, subinvestir às vésperas de um salto. Oportunidade: a janela 2026-2029, em que redesenho ainda é diferencial. Movimento: um processo core redesenhado de ponta a ponta por ano, com métrica de P&L, e nenhum piloto sem dono, critério de sucesso e data de morte.

Profissional de conhecimento. O que muda: a tarefa média morre antes do cargo; o valor migra para orquestração, julgamento e relação. Risco: ancorar identidade na tarefa. Oportunidade: o prêmio de quem supervisiona sistemas é real e crescente (payback mediano de agente em 5,1 meses significa que quem sabe implantá-los tem emprego em qualquer cenário, BCG e Forrester 2026). Movimento: construir portfólio de decisões assinadas, não de tarefas executadas.

Governo e regulador (lente Brasil). O que muda: o Brasil adota rápido (mercado pioneiro em uso, EY 2026; 42% das empresas com uso transformador contra 34% da média global, Deloitte 2026) e governa devagar (PL 2338 três anos em tramitação). Risco: repetir o padrão de importar regulação europeia com dez anos de defasagem e zero política industrial junto. Oportunidade: energia limpa mais escala digital mais dados públicos estruturados é uma combinação que pouquíssimos países têm. Movimento: aprovar o marco com regulamentação setorial ágil, e tratar computação soberana e eficiência energética como a mesma agenda.

Investidor. O que muda: o retorno migra da camada de modelo (margens em compressão em três dos quatro mundos de 2040) para infraestrutura energética, camada de aplicação com dado proprietário e os "picks and shovels" físicos. Risco: a assimetria capex-receita atual (US\$ 662 bi fora de balanço, Moody's 2026) sugere que parte do ciclo será corrigida; o timing é imprevisível, a direção não. Movimento: separar a tese "IA transforma a economia" (robusta) da tese "este ativo captura a transformação" (frágil), porque a primeira não paga a segunda.

Educação e família. O que muda: 80% dos estudantes americanos de ensino médio e superior já usam IA para tarefas escolares com só 6% dos professores relatando políticas claras (AI Index 2026), e 72% dos adolescentes usam IA como companhia (Common Sense Media 2025). O risco não é a cola, é a atrofia: terceirizar o pensar durante a década em que se aprende a pensar. Movimento: menos proibição de ferramenta, mais prática deliberada de decisão com consequência, que é a competência que todos os cenários premiam e nenhum currículo atual ensina.

14

Framework de monitoramento

O report vira sistema de alerta com dez indicadores, cada um amarrado ao cenário que sinaliza. Cadência sugerida: revisão trimestral leve, anual profunda.

INDICADOR	FRE-QUÊNCIA	LEITURA DE SINAL
Distância entre melhor modelo aberto e fechado (Arena e AI Index)	Trimes-tral	Crescendo por 12 meses: Oligopólio. Em zero por 18 meses: Eletricidade Cognitiva ou Tecnologia Normal.
Ganho geracional em benchmarks de agência (OSWorld, HLE, SWE-Bench)	Semes-tral	Dois ciclos de ganho marginal: quadrantes de platô. Saltos mantidos: quadrantes de salto.
Capex agregado dos hyperscalers, nominal e vs receita de IA	Trimes-tral	Acelerando com receita acompanhando: salto. Corte real por 2 anos: platô. Descolamento crescente: risco de correção (wild card).
Consumo elétrico de data centers vs faixa IEA (945 TWh em 2030)	Anual	Acima da faixa: pressão de Oligopólio (gargalo favorece grandes). Abaixo: Headwinds, platô.
Taxa de pilotos de agente em produção (S&P, McKinsey)	Semes-tral	Acima de 30% até 2028: Difusão com Atrito confirmada. Estagnada em ~12%: freio institucional mais forte que o previsto.
Regulação de acesso a computação (EUA, China, UE)	Trimes-tral	Licenciamento estatal de computação: Oligopólio. Foco migrando para responsabilidade por uso: Eletricidade Cognitiva.
Cortes de trabalho ligados a IA como % da força de trabalho	Anual	Abaixo de 1% ao ano: atrito. Acima de 2% por dois anos: transição dura, antecipa turbulência do H2.
Preço médio e volume de humanoides (Omdia, IFR)	Anual	Abaixo de US\$ 15 mil com >100 mil unidades/ano: IA ganhando corpo, acelera todos os cenários de salto.
Aprovação e regulamentação do PL 2338 no Brasil	Trimes-tral	Aprovado até 2027 com regulamentação setorial: Brasil adaptador competente. Travado além de 2028: adoção sem governança, risco local sobe.
Incidentes documentados de IA e res-posta regulatória (AI Index)	Anual	Crescimento com resposta institucional: Tecnologia Normal. Cresci-mento sem resposta: erosão de confiança, freia adoção de alto risco.

15

Folha de fontes

Transparência é parte do produto. Fornecedor nunca é fato neutro; imprensa carrega recorte; o tier está marcado para o leitor pesar cada dado.

FONTE	DOCUMENTO	DATA	TIER	USO NO REPORT
IEA	Energy and AI; Key Questions on Energy and AI	2025-2026	Institucional	Energia, gargalos físicos, projeções de demanda (Scanning, cenários, monitoramento)
Stanford HAI	AI Index Report 2026	2026	Institucional e acadêmica	Capacidade, adoção, incidentes, investimento, educação (Sumário, Scanning, Radar)
AISI / NCSC (Reino Unido)	Avaliação de capacidade cibernética de modelos de fronteira	2026	Primária e oficial	Wild card de ataque autônomo (Radar, cenário Oligopólio)
Conselho da UE; Covington; White & Case	Digital Omnibus / emendas ao AI Act	2026	Primária e oficial (com análise jurídica)	Regulação europeia (Scanning político, framing)
JPMorgan Research	Midyear Outlook 2026	2026	Fornecedor (banco, tese construtiva)	Capex US\$ 5,5 tri e US\$ 650 bi-1,1 tri (Sumário, Scanning econômico)
MIT NANDA	The GenAI Divide	2025	Institucional e acadêmica	95% de pilotos sem impacto; 67% vs 33% em builds (Sumário, implicações)
McKinsey	Global AI Survey	2025	Fornecedor (consultoria)	88% adoção, 39% EBIT, 2x com redesenho (Scanning, implicações)
Gartner; Forrester; BCG; S&P	Projeções de agentes, ROI, cancelamentos	2025-2026	Fornecedor	Adoção de agentes, payback 5,1 meses, 40% cancelados (Radar, implicações, com hedge)
Challenger, Gray & Christmas	Rastreamento de cortes nos EUA	2026	Institucional	IA como motivo líder de demissões (Sumário, Scanning)
Forbes Brasil; Gazeta Mercantil; Hardware.com.br	Cobertura de demissões por IA	2026	Imprensa	Números de cortes 2026, Goldman e projeções (Scanning social e econômico)
Goldman Sachs; PwC; Acemoglu (MIT); Penn Wharton	Projeções de impacto no PIB	2023-2025	Fornecedor e acadêmica (intervalo entre elas usado como dado)	Faixa de impacto macroeconômico (Scanning econômico)
CSIS; CNAS; CFR; Mayer Brown	Análises de controles de exportação e localização chinesa	2025-2026	Institucional	Bifurcação de chips, H200, EUV (Scanning político, cenários)
TeamDay; Kili Technology	Panorama de modelos e benchmarks 2026	2026	Imprensa especializada	Lançamentos de fronteira, saturação de benchmark (Scanning tecnológico)

Fonte	Documento	Data	Tier	Uso no Report
Common Sense Media; APA; OpenAI-MIT Media Lab	Estudos de IA e companhia	2025-2026	Institucional e acadêmica	IA íntima, teens, solidão (Scanning social, Radar)
Narayanan e Kapoor (Princeton)	AI as Normal Technology	2025	Acadêmica	Tese de difusão lenta (CLA, cenário Tecnologia Normal)
International AI Safety Report	Edição 2026	2026	Institucional	Comportamento diferencial teste vs produção (Scanning valores, Radar)
KraneShares; Omdia; Rest of World	Dados de robótica humanoide	2026	Fornecedor e imprensa	Volumes, preços, BMW-Figure (Scanning tecnológico, Radar, com hedge)
Moody's (via análises de mercado)	Compromissos de leasing fora de balanço	2026	Institucional (citação indireta)	Wild card de correção financeira (Radar, implicações)
Deloitte; EY Brasil; Abiacom	State of AI in the Enterprise; rankings de adoção BR	2026	Fornecedor	Posição brasileira de adoção (Scanning, implicações)
Senado Federal; gov.br; Barbieri Advogados	PL 2338, PBIA, ANPD	2024-2026	Primária e oficial (com análise jurídica)	Marco regulatório brasileiro (Scanning político, monitoramento)
FutureSearch; AI Impacts; aiMultiple	Agregações de previsões de AGI	2023-2026	Institucional e imprensa especializada	Distribuição de timelines, vies de horizonte móvel (Radar, eixos)