

ARTIGO ACADÊMICO

A privatização do limite necessário

Economia de aperfeiçoamento, capacidade algorítmica e
soberania em um planeta de recursos finitos

*The privatisation of the necessary limit: economy of improvement, algorithmic
capacity, and sovereignty on a planet of finite resources*

Jandislei Antonio Genova

*Advogado, autor e pesquisador independente em Direito, inteligência
artificial e governança institucional*

Julho de 2026 — segunda versão revista e ampliada

Versão integral para circulação acadêmica e desenvolvimento de pesquisa

Resumo

O debate sobre a sustentabilidade da inteligência artificial costuma concentrar-se na eficiência energética, na redução do custo computacional ou na seleção de modelos menores. Essas contribuições são indispensáveis, mas não resolvem uma questão anterior: quem possui autoridade para definir quanta capacidade algorítmica será suficiente, para quais finalidades, a que custo e segundo quais critérios distributivos? Este artigo investiga a passagem de um problema técnico de otimização para um problema jurídico e político de poder. Adota pesquisa qualitativa, bibliográfica e documental, com revisão narrativa crítica submetida a protocolo explícito de busca, seleção e diferenciação conceitual, além de análise jurídico-dogmática e aplicação ilustrativa a um caso-tipo de contratação pública. A literatura foi organizada em quatro blocos: sustentabilidade e ciclo de vida; seleção e roteamento de modelos; governança e desigualdade computacional; justiça distributiva e soberania digital. Nesse diálogo, cinco expressões são propostas como categorias analíticas, e não como reivindicação de criação ex nihilo: economia de aperfeiçoamento, economia política da capacidade algorítmica, privatização do limiar de suficiência, suficiência algorítmica distributiva e soberania sobre o limiar. Sustenta-se que o escalonamento da IA constitui trajetória sociotécnica orientada por capital, infraestrutura, métricas e decisões institucionais. A eficiência pode reduzir o custo unitário e, simultaneamente, ampliar a demanda agregada, reforçar dependências e permitir que fornecedores administrem níveis desiguais de capacidade. Como especialização sociotécnica e distributiva da proporcionalidade, propõe-se o Teste do Limiar de Suficiência Algorítmica (TLSA): nem mais capacidade do que a finalidade justifica; nem menos capacidade do que o risco e os direitos exigem. O instrumento combina piso de direitos, justificação do incremento, alternativa menos intensiva, análise de ciclo de vida, distribuição de ônus, possibilidade de saída e revisão. Conclui-se que a soberania é capturada quando a sociedade suporta os custos da IA, mas não participa da definição de sua escala, finalidade, acesso e limites.

Palavras-chave: inteligência artificial; economia de aperfeiçoamento; capacidade algorítmica; suficiência algorítmica distributiva; influência cognitiva estrutural; soberania digital.

Abstract

The sustainability debate surrounding artificial intelligence often focuses on energy efficiency, lower computational costs, or the selection of smaller models. These contributions are indispensable, yet they do not answer a prior question: who has the authority to determine how much algorithmic capacity is sufficient, for which purposes, at what cost, and under which distributive criteria? This article examines the transition from a technical optimisation problem to a legal and political problem of power. It adopts qualitative bibliographic and documentary research, combining a critical narrative review under an explicit search and selection protocol, doctrinal legal analysis, and an illustrative application to a public-procurement case type. The literature is organised into four clusters: sustainability and life cycle; model selection and routing; compute governance and inequality; and distributive justice and digital sovereignty. In dialogue with these fields, five expressions are proposed as analytical categories rather than as claims of creation ex nihilo: economy of improvement, political economy of algorithmic capacity, privatisation of the sufficiency threshold, distributive algorithmic sufficiency, and sovereignty over the threshold. AI scaling is understood as a sociotechnical trajectory shaped by capital, infrastructure, metrics, and institutional decisions. Efficiency may reduce unit costs while increasing aggregate demand, reinforcing dependencies, and enabling providers to administer unequal levels of capacity. As a sociotechnical and distributive specification of proportionality, the article proposes the Algorithmic Sufficiency Threshold Test (ASTT): no more capacity than the purpose justifies; no less capacity than risk and rights require. The instrument combines a rights floor, justification of marginal improvement, less resource-intensive alternatives, life-cycle assessment, distribution of burdens, exit capacity, and review. Sovereignty is captured when society bears AI's costs but does not participate in defining its scale, purpose, access, and limits.

Keywords: artificial intelligence; economy of improvement; algorithmic capacity; distributive algorithmic sufficiency; structural cognitive influence; digital sovereignty.

1 INTRODUÇÃO

O acesso cotidiano a sistemas de inteligência artificial produz uma ilusão de imaterialidade. Uma pergunta é digitada, uma resposta aparece e toda a cadeia que tornou possível essa interação permanece fora do campo perceptivo: extração mineral, fabricação de semicondutores, consumo de água, geração e transmissão de energia, construção de centros de dados, redes de telecomunicações, trabalho humano de classificação e moderação, aquisição de bases de dados, desenvolvimento de modelos, serviços de nuvem e contratos que distribuem riscos, propriedade e controle. A interface faz parecer instantâneo aquilo que é materialmente extenso, juridicamente organizado e politicamente disputado.

Essa invisibilidade favorece duas reduções. A primeira trata a sustentabilidade da IA como problema exclusivamente técnico: bastaria produzir chips mais eficientes, comprimir modelos, reduzir parâmetros ou encaminhar cada solicitação ao sistema de menor custo capaz de respondê-la. A segunda converte a escassez em justificativa para uma hierarquia entre pessoas: sistemas superiores seriam reservados a pesquisadores, governos ou profissionais considerados produtivos, enquanto à maioria caberiam versões limitadas. A primeira redução ignora relações de poder; a segunda transforma uma preocupação ambiental legítima em arquitetura de segregação cognitiva.

Este artigo rejeita ambas. Recursos planetários são finitos, mas pessoas não podem ser classificadas por uma autoridade privada ou estatal segundo um suposto merecimento de inteligência. A questão normativa não é decidir quem “merece” uma IA avançada. É determinar quando um incremento de capacidade gera benefício marginal demonstrável, quem decide sobre sua produção e distribuição, como os custos são internalizados e quais pisos de acesso decorrem de direitos, riscos e finalidades públicas.

O problema de pesquisa pode ser formulado da seguinte maneira: **como estabelecer limites materialmente sustentáveis para o desenvolvimento e o uso da IA sem converter a suficiência em instrumento de estratificação, censura, subordinação tecnológica ou concentração da autoridade cognitiva?**

A hipótese é que a eficiência técnica e o *right-sizing* são necessários, mas insuficientes. Quando os mesmos agentes controlam modelos, nuvem, interfaces, métricas, preços e registros de uso, eles podem definir unilateralmente o que será considerado suficiente para cada usuário e instituição. Nesse cenário, a alocação eficiente de recursos também pode funcionar como distribuição privada de oportunidades cognitivas. O problema deixa de ser apenas o custo de uma inferência e passa a incluir o poder de determinar a qualidade, a

profundidade, o tempo de processamento, os instrumentos disponíveis, a possibilidade de auditoria e a continuidade do serviço.

O objetivo geral é construir uma arquitetura conceitual e normativa para governar o limiar de suficiência da IA. Os objetivos específicos são: a) diferenciar eficiência, frugalidade, suficiência digital, *right-sizing* e divisão computacional; b) explicar por que o escalonamento da IA constitui trajetória dirigida, e não evolução natural; c) formular as categorias economia de aperfeiçoamento, economia política da capacidade algorítmica, privatização do limiar de suficiência, suficiência algorítmica distributiva e soberania sobre o limiar; d) relacionar a distribuição de capacidade à influência cognitiva estrutural; e) propor um teste normativo aplicável ao desenvolvimento, à contratação e ao uso institucional de sistemas; e f) indicar mecanismos de governança compatíveis com sustentabilidade, igualdade, concorrência e autonomia.

1.1 Delineamento, protocolo de busca e método jurídico

Trata-se de pesquisa qualitativa, interdisciplinar, bibliográfica e documental, desenvolvida mediante revisão narrativa crítica, reconstrução conceitual e análise jurídico-dogmática. A revisão narrativa foi escolhida porque o problema atravessa campos que utilizam objetos, vocabulários e métodos distintos — engenharia de aprendizagem de máquina, estudos ambientais, economia política, filosofia da justiça e Direito — e porque o objetivo não é estimar um único efeito, mas testar se esses campos, quando relacionados, tornam visível uma autoridade ainda tratada de maneira fragmentária. Essa opção não dispensa transparência metodológica.¹

As buscas foram realizadas e atualizadas entre 23 e 24 de julho de 2026 em repositórios acadêmicos e editoriais de acesso público, incluindo arXiv, ACM Digital Library, AAAI/ACM AIES, ACL Anthology, JMLR, SAGE, Nature e Oxford Academic, com rastreamento regressivo das referências centrais. Foram utilizados, isoladamente e em combinações, os descritores em inglês “AI compute governance”, “compute divide”, “public compute”, “computing commons”, “Green AI”, “Frugal AI”, “digital sufficiency”, “LLM right-sizing”, “LLM routing”, “AI environmental lifecycle”, “digital sovereignty”, “sufficientarianism” e “threshold justice”; e, em português, “capacidade computacional de IA”, “infraestrutura pública de IA”, “soberania digital”, “suficiência algorítmica” e “regulação ambiental da IA”. Para o recorte normativo brasileiro e europeu, foram consultadas as bases

¹A revisão é narrativa e crítica. As expressões propostas não são apresentadas como prova de prioridade terminológica absoluta, mas como categorias empregadas em sentido específico e delimitado. Pesquisas futuras deverão testar sua utilidade explicativa e normativa.

oficiais do Planalto, Câmara dos Deputados, Senado Federal, CNJ, MCTI, MDIC, MCom, LNCC, EUR-Lex, UNESCO, OCDE, IEA e UNEP.

Foram incluídos: a) artigos revisados por pares diretamente relacionados aos quatro blocos analíticos; b) pré-publicações apenas quando representavam evidência técnica emergente não substituída por publicação consolidada, sempre qualificadas como tal; c) livros de referência sobre economia política, infraestrutura e soberania digital; d) normas, projetos legislativos e atos administrativos em fonte oficial; e e) relatórios institucionais que apresentassem método ou dados identificáveis. Foram excluídos textos meramente promocionais, notícias sem documento primário correspondente, duplicatas e materiais cuja contribuição fosse apenas tangencial. Fontes jornalísticas foram usadas somente para a gênese pública do problema e não sustentam conclusões técnicas ou normativas.

No plano jurídico, o trabalho distingue três operações. A primeira identifica fundamentos de direito positivo (*de lege lata*) capazes de incidir sobre decisões de capacidade: Constituição, proteção ambiental, defesa da concorrência e do consumidor, proteção de dados, contratação pública e governança setorial. A segunda realiza integração funcional entre esses regimes sem ampliar artificialmente o âmbito de incidência de cada norma. A terceira formula propostas institucionais (*de lege ferenda*), explicitamente apresentadas como recomendações e não como deveres já positivados. A proporcionalidade é utilizada como estrutura de comparação; o TLISA é examinado como sua especialização sociotécnica e distributiva, não como substituto autônomo do raciocínio jurídico.

Quadro 1 — Protocolo de constituição e uso do corpus

Bloco	Fontes prioritárias	Pergunta de inclusão	Função no argumento
Sustentabilidade e ciclo de vida	estudos empíricos, revisões técnicas, IEA e UNEP	o trabalho mede ou diferencia energia, água, carbono, hardware, treinamento e inferência?	estabelecer materialidade, incerteza e efeito rebote
Seleção e roteamento	artigos sobre <i>right-sizing</i> , cascatas e roteadores de modelos	há comparação entre qualidade, custo e capacidade em tarefas identificáveis?	mostrar viabilidade técnica e riscos de alocação opaca
Governança e desigualdade computacional	estudos sobre <i>compute governance</i> , concentração, nuvem e computação pública	a fonte explica propriedade, acesso, geografia ou poder regulatório da infraestrutura?	localizar quem controla o limiar e a possibilidade de saída
Justiça, Direito e soberania	filosofia da suficiência, normas e doutrina jurídica	a fonte oferece critério para piso, teto, proporcionalidade, direitos ou competência institucional?	construir e limitar normativamente o TLISA

Fonte: elaboração do autor.

O corpus final reúne 66 referências e é integralmente identificável ao término do artigo. Como a investigação não foi originalmente desenhada como revisão sistemática, os

resultados brutos de cada mecanismo de busca não foram exportados em registro imutável; além disso, páginas oficiais e repositórios dinâmicos alteram seus resultados. A transparência alcançada é, portanto, documental e lógica, não equivalente a protocolo PRISMA. A estratégia não autoriza inferências quantitativas sobre prevalência, magnitude ou causalidade e tampouco estima a pegada de um modelo específico. Sua validade depende da coerência da reconstrução, da correspondência entre proposições e fontes e da possibilidade de refutação futura por estudos empíricos, comparação institucional e aplicação do teste a decisões reais.

A contribuição reside menos em afirmar que a IA consome recursos — fato amplamente estudado — e mais em deslocar a pergunta para a autoridade que define o necessário. Green AI, IA frugal, suficiência digital, roteamento e *right-sizing* oferecem respostas relevantes à relação entre desempenho e custo. A governança da computação e a divisão computacional explicam a desigualdade no acesso à infraestrutura e seu uso como ponto de intervenção regulatória. A filosofia da suficiência discute o conteúdo e a legitimidade dos limiares. A computação pública procura criar alternativas aos provedores dominantes. O argumento proposto é uma síntese normativa original desses campos: conecta incremento de capacidade, ciclo de vida, formação da vontade, distribuição de possibilidades decisórias, possibilidade de saída e captura da soberania.

O artigo também se distingue da formulação “Quanto mais IA, menos IA”, anteriormente utilizada pelo autor para examinar como a expansão indiscriminada de sistemas pode reduzir qualidade, confiança e autonomia. O objeto atual não é a saturação quantitativa da IA, mas a governança do incremento marginal de capacidade e do poder de estabelecer seu limiar.²

A tese pode ser antecipada em duas proposições:

O princípio da suficiência não deve limitar pessoas; deve limitar o desperdício, a expansão injustificada e o poder privado de distribuir capacidade cognitiva segundo critérios opacos.

A soberania é capturada quando a sociedade suporta os custos materiais da inteligência artificial, mas permanece sem poder para decidir sua finalidade, escala, critérios de acesso e limites.

2 A INFRAESTRUTURA QUE A INTERFACE OCULTA

2.1 Da resposta instantânea à cadeia sociotécnica

Inteligência artificial não é apenas modelo. Um sistema em operação depende de uma cadeia que inclui minerais, equipamentos de fabricação, semicondutores, redes, energia,

²A referência anterior é: GENOVA, Jandislei Antonio. Quanto mais IA, menos IA. LinkedIn, 2026. Disponível em: <https://www.linkedin.com/feed/update/urn:li:activity:7481343941032898560>. Acesso em: 24 jul. 2026.

refrigeração, centros de dados, nuvem, software, dados, trabalho, propriedade intelectual, capital e regras de acesso. Crawford (2021) descreve a IA como indústria extrativa, cuja aparência abstrata repousa sobre territórios, recursos e trabalho. A análise de ciclo de vida do Programa das Nações Unidas para o Meio Ambiente reforça que a avaliação ambiental não pode começar no treinamento e terminar na inferência: deve alcançar extração, fabricação, transporte, operação e descarte (UNEP, 2024).

A materialidade não autoriza generalizações imprecisas. “Consumo da IA” não é uma unidade homogênea. Treinamento, ajuste, inferência, armazenamento, transmissão, resfriamento e fabricação possuem perfis distintos; matrizes energéticas e condições climáticas alteram resultados; fornecedores divulgam informações incompletas; melhorias algorítmicas mudam rapidamente as relações entre custo e desempenho. Trabalhos sobre energia e carbono em aprendizagem de máquina demonstram a importância de métricas e relatórios comparáveis (STRUBELL; GANESH; MCCALLUM, 2019; HENDERSON et al., 2020). Luccioni, Jernite e Strubell (2024) mostram, em comparação sistemática de tarefas, que arquiteturas generativas de propósito geral podem demandar muito mais energia em inferência do que sistemas especializados. A revisão de Chien et al. (2026) acrescenta duas cautelas decisivas: emissões incorporadas na fabricação e construção podem representar parcela majoritária do impacto de grandes centros orientados à IA; e medidas de redução de água podem produzir compensações indesejadas em carbono. A incerteza, portanto, não elimina o dever de governança. Ela impede a redução do problema a uma única métrica.

Segundo a Agência Internacional de Energia, centros de dados responderam por aproximadamente 1,5% do consumo mundial de eletricidade em 2024. A IA representa apenas uma parcela das cargas desses centros, razão pela qual o número não deve ser atribuído integralmente a ela. Ainda assim, a agência registra que um centro de dados convencional pode demandar entre 10 e 25 megawatts, enquanto instalações de hiperescala orientadas à IA podem superar 100 megawatts e se concentram espacialmente, pressionando redes locais (IEA, 2025). Em atualização publicada em abril de 2026, a mesma agência estimou que o consumo elétrico dos centros de dados cresceu 17% em 2025, enquanto instalações concentradas em IA cresceram mais rapidamente, apesar da forte queda do consumo por tarefa (IEA, 2026). A distinção é decisiva: uma participação global aparentemente limitada pode produzir efeitos territoriais intensos sobre energia, água, preço, licenciamento e prioridade de conexão.

O impacto também é distributivo. A infraestrutura pode ser instalada em território que fornece água, energia, incentivos fiscais e estabilidade institucional, enquanto propriedade,

renda, modelos e poder decisório permanecem em outro lugar. Benefícios e ônus não acompanham necessariamente a mesma geografia. Em economias periféricas, a promessa de integração à cadeia de IA pode coexistir com dependência de chips, nuvens, modelos, padrões e capital externos.

2.2 Eficiência não garante redução absoluta

Eficiência mede a relação entre resultado e recurso. Um modelo que produz resposta semelhante com menos energia, menos memória ou menor latência é tecnicamente preferível, mantidas as demais condições. O problema surge quando a redução do custo unitário barateia e expande o uso. Mais consultas, contextos maiores, respostas mais longas, agentes em funcionamento contínuo e IA incorporada a objetos antes não automatizados podem neutralizar ou superar a economia obtida por operação.

Esse mecanismo é conhecido, em termos gerais, como efeito rebote. A suficiência digital foi desenvolvida precisamente porque eficiência e consistência produtiva não asseguram diminuição absoluta dos impactos das tecnologias de informação. Santarius et al. (2023) identificam dimensões de suficiência de hardware, software, usuário e economia, conectando a digitalização a limites planetários e aos efeitos de indução de demanda. A atualização da IEA (2026) oferece uma expressão contemporânea do paradoxo: tarefas simples tornaram-se muito mais eficientes, mas raciocínio prolongado, geração de vídeo e sistemas agênticos podem consumir centenas ou milhares de vezes mais energia por consulta, ao mesmo tempo que a base de usuários e a frequência de uso se expandem.

Na IA, o rebote adquire forma adicional: a corrida por desempenho cria novos padrões de normalidade. Uma capacidade que ontem era excepcional torna-se requisito mínimo de mercado; aplicações são redesenhadas para consumi-la; modelos menores passam a parecer inadequados não porque falharam em uma finalidade, mas porque a referência foi deslocada. O ganho de eficiência alimenta a expansão da fronteira e a fronteira redefine o suficiente.

2.3 Limites planetários e limites institucionais

A noção de limites planetários oferece referência para compreender que sistemas econômicos e tecnológicos operam em um espaço biofísico finito (ROCKSTRÖM et al., 2009; RICHARDSON et al., 2023). Aplicá-la à IA não significa estabelecer uma cota mundial simples de operações. Água, energia, carbono, minerais, biodiversidade e resíduos não são plenamente comensuráveis; impactos variam por território, tempo e matriz produtiva. Um único número poderia ocultar desigualdades e incentivar compensações artificiais.

O limite relevante é, portanto, simultaneamente material e institucional. Material, porque a capacidade computacional exige recursos. Institucional, porque regras, contratos, mercados e políticas determinam quais projetos serão construídos, conectados e financiados. A escassez computacional não chega à sociedade em estado puro: ela já nasce politicamente administrada.

3 O QUE A LITERATURA JÁ EXPLICA — E O QUE PERMANECE EM ABERTO

3.1 Green AI

Schwartz et al. (2020) formularam a agenda Green AI como contraponto a uma cultura de pesquisa concentrada exclusivamente em acurácia. A proposta defende que resultados sejam acompanhados de custo computacional e que eficiência seja tratada como critério de mérito científico. A contribuição é fundamental: torna visível o preço material de ganhos marginais e questiona benchmarks que premiam desempenho sem contabilizar recursos.

Green AI, entretanto, não resolve por si só a distribuição da autoridade. Mesmo quando custos são medidos, permanece a pergunta sobre quem escolhe quais ganhos justificam esses custos, quem possui infraestrutura para persegui-los e quem suporta as externalidades. Henderson et al. (2020), Luccioni, Jernite e Strubell (2024), UNEP (2024) e Chien et al. (2026) tornam o relato ambiental mais preciso, mas a mensuração, isoladamente, não oferece regra de prioridade, competência ou contestação. Ela informa a decisão; não substitui o procedimento político e jurídico.

3.2 Inteligência artificial frugal

IA frugal reúne técnicas e princípios para desenvolver soluções úteis sob restrições de energia, dados, hardware, conectividade e custo. Compressão, quantização, poda, destilação, otimização de hardware, aprendizagem eficiente e adequação ao contexto integram esse campo. Arga et al. (2025) organizam a literatura como agenda de inovação sustentável e acessível em condições de recursos limitados.

Frugalidade não equivale a precariedade. Um sistema local, menor e especializado pode oferecer melhor privacidade, latência, auditabilidade e controle do que um modelo geral de fronteira. Todavia, a expressão também pode ser apropriada para legitimar tecnologia inferior destinada a populações ou instituições com menor poder de compra. Sem um piso de direitos, desempenho e recurso, “frugal” pode tornar-se eufemismo para serviço de segunda classe. A contribuição deste artigo começa precisamente onde a frugalidade deixa de ser

decisão interna de engenharia e passa a definir quais capacidades estarão disponíveis a sujeitos e instituições.

3.3 Suficiência digital

A suficiência digital desloca o debate da eficiência relativa para o impacto absoluto. Santarius et al. (2023) perguntam não apenas como realizar a mesma tarefa com menos recursos, mas quais quantidades de hardware, tráfego, consumo e crescimento são compatíveis com uma transformação ambiental efetiva. A abordagem inclui políticas e escolhas econômicas, não apenas engenharia.

O presente artigo herda esse deslocamento, mas modifica seu objeto. Em vez de examinar o consumo digital em geral, concentra-se na capacidade algorítmica disponibilizada para conhecer, interpretar, prever, recomendar e decidir. O que está em disputa não é apenas volume de dados ou energia: é a distribuição de meios cognitivos socialmente relevantes. Essa mudança de objeto exige relacionar suficiência ambiental a um piso de capacidades, pois reduzir impacto absoluto não legitima reduzir proteção, segurança ou acesso abaixo do necessário ao exercício de direitos.

3.4 Right-sizing e seleção orientada à tarefa

O *right-sizing* procura selecionar o modelo cuja capacidade seja adequada à tarefa, evitando o uso automático do sistema mais amplo. Uma pré-publicação experimental indica que modelos menores podem alcançar resultados satisfatórios em atividades organizacionais específicas, com menor custo e maior possibilidade de implantação local (HAASE et al., 2025). Seus resultados não autorizam generalização para tarefas críticas, mas o mérito da abordagem consiste em substituir a maximização abstrata de benchmarks por avaliação contextual.³

A literatura de roteamento confirma que essa escolha pode ser tecnicamente estruturada. FrugalGPT utiliza adaptação, aproximação e cascatas de modelos para reduzir custo sem necessariamente reduzir desempenho (CHEN; ZAHARIA; ZOU, 2023). O Hybrid LLM encaminha solicitações entre modelos segundo dificuldade prevista e nível de qualidade desejado (DING et al., 2024). O RouteLLM aprende a rotear a partir de dados de preferência, equilibrando custo e qualidade (ONG et al., 2024). Esses resultados demonstram viabilidade, não legitimidade distributiva: custo e acurácia agregada não capturam direitos, explicabilidade, dependência, erro assimétrico nem importância institucional da tarefa.

³Haase et al. (2025) é uma pré-publicação (working paper) e emprega conjunto limitado de tarefas e modelos. É citada como evidência emergente da abordagem, não como demonstração definitiva de superioridade geral de modelos menores.

Se aplicado sem garantias, o encaminhamento pode reproduzir o problema que pretende resolver. O fornecedor define categorias de tarefa, níveis de qualidade, limites de contexto, prioridade, tempo e acesso a ferramentas. O usuário pode desconhecer que recebeu sistema reduzido, não saber quais capacidades foram suprimidas e não possuir meio de contestar a classificação. O *right-sizing* técnico pode tornar-se *downscaling* comercial opaco. A lacuna não é de engenharia do roteador, mas de governança da decisão de rotear.

3.5 Divisão computacional e concentração da pesquisa

Ahmed e Wahed (2020) denominam *compute divide* a desigualdade de acesso à capacidade computacional que favorece grandes empresas e universidades de elite. Ahmed, Wahed e Thompson (2023) mostram que a indústria passou a dominar insumos centrais da pesquisa contemporânea em IA: computação, dados e talento. A consequência não é apenas vantagem mercadológica; é capacidade de definir problemas de pesquisa, benchmarks, arquiteturas e futuros tecnológicos, enquanto alternativas orientadas ao interesse público se tornam escassas.

Sastry et al. (2024) tratam a computação como ponto de intervenção regulatória porque é relativamente detectável, quantificável e excludente e porque sua cadeia é concentrada. O mesmo estudo adverte que políticas ingênuas de governança da computação podem intensificar vigilância, centralização e efeitos econômicos adversos. Lehdonvirta, Wú e Hawkins (2024), ao mapear regiões de nuvem com GPU, distinguem *Compute North*, *Compute South* e *Compute Desert*: a geografia da infraestrutura condiciona não apenas quem desenvolve modelos, mas quais Estados possuem visibilidade e capacidade prática de impor regras. A desigualdade, portanto, alcança o próprio poder regulatório.

Relatórios do AI Now Institute observam que a computação constitui dependência central em pontos concentrados da cadeia, produzindo bloqueio, poder de mercado e controle sobre quem pode construir IA e com quais finalidades (VIPRA; MYERS WEST, 2023). A OCDE trata capacidade nacional como combinação de disponibilidade, efetividade e resiliência, incluindo sustentabilidade e soberania (OECD, 2023). Luitse e Denkena (2021), por sua vez, demonstram como grandes modelos distribuídos por APIs podem funcionar como componentes infraestruturais rentistas e aprofundar dependência. Esses trabalhos oferecem os antecedentes mais próximos da “economia política da capacidade algorítmica”; a diferença proposta aqui é dirigir a análise ao poder de fixar o nível funcional efetivamente entregue e contestável, não apenas ao acesso bruto à computação.

3.6 Computação pública e bens comuns computacionais

Propostas de *public compute* e *computing commons* respondem à concentração mediante infraestrutura financiada, orientada ou governada publicamente. O objetivo é permitir pesquisa, inovação e serviços de interesse coletivo fora da dependência integral de provedores dominantes. Davies e Vipra (2025) distinguem propriedade formal de orientação pública: infraestrutura estatal pode reproduzir concentração, vigilância ou dano ambiental se não houver governança, participação e finalidade pública.

Essa advertência impede uma resposta simplista. Transferir a administração do limiar de uma empresa para um órgão estatal não garante justiça. Vales para nuvens privadas podem ampliar acesso imediato e, ao mesmo tempo, subsidiar incumbentes; propriedade pública sem equipe, software portátil e regras de acesso pode conservar dependência; infraestrutura de fronteira pode absorver recursos que atenderiam melhor a numerosos projetos menores. A autoridade pública também pode discriminar, censurar, desperdiçar ou favorecer grupos. O problema é desenhar instituições contestáveis, transparentes e plurais, com direitos de acesso, revisão e saída.

3.7 Suficiência, limiares e justiça distributiva

Na filosofia política, o suficientarismo atribui importância moral a assegurar que pessoas ultrapassem um limiar relevante, em vez de tratar toda desigualdade como igualmente prioritária. Frankfurt (1987) formulou uma referência clássica; Fourie e Rid (2016) organizaram o debate aplicado; Axelsen e Nielsen (2016) sustentam que um limiar defensável precisa evitar ambiguidade, arbitrariedade, baixa exigência e cegueira a outros valores. Axelsen e Shields (2025) acrescentam que limiares podem funcionar melhor como demarcação do campo da justiça do que como solução automática para todas as suas exigências.

Essas objeções incidem diretamente sobre a proposta. “Suficiente” não pode significar qualquer nível que o fornecedor esteja disposto a entregar, nem um mínimo uniforme insensível ao risco. A moeda relevante também precisa ser definida: aqui, não é bem-estar, renda ou número de parâmetros, mas capacidade funcional situada — qualidade, contexto, ferramentas, auditabilidade, segurança, continuidade e possibilidade de contestação necessárias a determinada finalidade. O limiar inferior protege direitos e capacidades institucionais; o limiar superior não proíbe excelência, mas desloca para o proponente do incremento o ônus de justificar custos e externalidades adicionais.

Há ainda uma ambiguidade terminológica a evitar. Em aprendizagem de máquina, “suficiência” também designa uma família de critérios estatísticos de equidade ligada à paridade preditiva ou à calibração entre grupos (DICICCIO et al., 2023; BAUMANN; HANNÁK; HEITZ, 2022). **Suficiência algorítmica distributiva**, neste artigo, não é métrica de calibração. É princípio normativo de distribuição de capacidade sociotécnica. Uma implementação poderá empregar métricas de equidade, mas não se reduz a elas.

3.8 Matriz de diferenciação conceitual

Quadro 2 — Delimitação entre campos consolidados e categorias propostas

Campo ou categoria	Pergunta central	Contribuição principal	Limite diante do problema deste artigo
Green AI	Quanto desempenho é obtido por recurso consumido?	Mensuração e valorização da eficiência	Não define quem decide se o ganho marginal justifica o custo
IA frugal	Como produzir utilidade sob restrição de recursos?	Técnicas e desenho orientados à escassez e acessibilidade	Pode legitimar serviço inferior sem piso de direitos
Suficiência digital	Quanto consumo digital é compatível com limites planetários?	Redução absoluta e enfrentamento do rebote	Não se concentra na distribuição de capacidade decisória
Right-sizing	Qual modelo é adequado a determinada tarefa?	Seleção contextual e redução do superdimensionamento	Pode ser controlado de forma opaca pelo fornecedor
Roteamento e cascatas de modelos	Como equilibrar custo e qualidade entre sistemas heterogêneos?	Demonstra viabilidade técnica de encaminhamento dinâmico	Otimiza funções escolhidas pelo projetista e não resolve legitimidade, direitos ou recurso
Governança da computação	Como usar a computação como ponto de visibilidade, alocação e controle?	Identifica instrumento regulatório relativamente quantificável e concentrado	Pode ampliar vigilância e centralização; não define por si só o piso distributivo
Compute divide	Quem possui acesso à computação necessária para pesquisar e desenvolver IA?	Diagnóstico da desigualdade estrutural de infraestrutura	Descreve acesso, mas não oferece critério completo para o limiar
Public compute / computing commons	Como criar capacidade orientada ao interesse público?	Alternativa infraestrutural à concentração privada	Infraestrutura pública também requer regras de finalidade e acesso
Suficientarismo e justiça dos limiares	O que torna um patamar suficiente e moralmente relevante?	Oferece gramática de piso, prioridade e crítica à arbitrariedade	Não foi concebido para a cadeia material e funcional dos sistemas de IA
Influência cognitiva estrutural	Como arquiteturas informacionais condicionam etapas anteriores à decisão?	Explica a formação estrutural do campo decisório	Precisa incorporar a camada material e distributiva da capacidade
Economia de aperfeiçoamento	Quando o incremento marginal de capacidade é justificável?	Disciplina a expansão e exige demonstração de benefício adicional	Depende de procedimento legítimo e indicadores não redutíveis a preço
Privatização do limiar de suficiência	Quem define o suficiente e com qual poder de contestação?	Torna visível a autoridade embutida na alocação técnica	Requer salvaguardas contra captura pública e privada

Campo ou categoria	Pergunta central	Contribuição principal	Limite diante do problema deste artigo
Suficiência algorítmica distributiva	Qual piso e qual teto protegem simultaneamente direitos e limites?	Une sustentabilidade, igualdade e não discriminação	Exige aplicação contextual e revisão periódica

Fonte: elaboração do autor, com base na literatura examinada.

A tabela não reivindica a inexistência de aproximações entre os campos. Ao contrário, explicita suas interseções para evitar que uma nova expressão apenas renomeie conhecimento consolidado. A originalidade defendida é de **síntese e especialização**: os termos propostos organizam, em uma mesma arquitetura, incremento marginal, ciclo de vida, piso de direitos, poder de alocação, dependência e possibilidade de saída. São proposições analíticas abertas à refutação, não territórios acadêmicos autônomos já estabelecidos nem reivindicações de prioridade terminológica absoluta.

4 A EVOLUÇÃO TECNOLÓGICA DIRIGIDA

4.1 Por que “evolução natural da IA” é uma formulação enganosa

Modelos maiores e mais capazes são frequentemente descritos como etapa natural do progresso. A metáfora sugere que o desenvolvimento seguiria lógica própria, comparável a processo biológico inevitável, cabendo à sociedade adaptar-se. Essa linguagem oculta escolhas: quais projetos recebem capital; quais métricas definem sucesso; quais infraestruturas recebem energia e licenças; quais riscos são tolerados; quais alternativas deixam de ser financiadas.

Winner (1980) demonstrou que artefatos e sistemas técnicos podem incorporar relações políticas. Feenberg (1999) sustenta que a tecnologia não constitui destino autônomo, mas campo de disputa social. Acemoglu e Johnson (2023) mostram historicamente que a direção da inovação e a distribuição de seus ganhos dependem de instituições, poder e escolhas sobre automação e complementaridade humana. Na IA, a escala tornou-se sinal de liderança porque mercados financeiros, laboratórios, governos e benchmarks passaram a recompensá-la.

Isso não significa que ganhos de escala sejam fictícios. Em determinadas tarefas, mais dados, parâmetros, computação e ferramentas produzem capacidades relevantes. A crítica dirige-se à passagem indevida do enunciado empírico — “mais recursos melhoraram este resultado” — para o imperativo político — “toda expansão disponível deve ser realizada”. O possível não cria automaticamente o necessário.

4.2 A fronteira como produto institucional

A “fronteira” é definida por competições, avaliações e anúncios construídos por agentes com capacidade de participar. Se universidades, empresas menores, instituições do Sul Global e organizações públicas carecem de computação, seus problemas e métodos aparecem menos na literatura e nos benchmarks. A desigualdade de insumos transforma-se em desigualdade de agenda.

O domínio industrial descrito por Ahmed, Wahed e Thompson (2023) permite que empresas não apenas alcancem a fronteira, mas influenciem o significado de avanço. Produtos gerais, escaláveis e integráveis a plataformas globais recebem maior visibilidade do que sistemas locais, verificáveis, modestos e orientados a necessidades públicas. A medida de excelência pode privilegiar aquilo que o detentor da infraestrutura sabe produzir e monetizar.

4.3 O aperfeiçoamento marginal e seus custos invisíveis

À medida que o desempenho cresce, cada incremento pode exigir recursos desproporcionais. Mesmo quando a relação não é monotônica, o último ponto percentual de um benchmark pode custar mais energia, capital e concentração do que ganhos anteriores. Além do custo ambiental, existe custo de oportunidade: computação, pesquisadores, energia e financiamento destinados à corrida de fronteira deixam de atender outras necessidades científicas e sociais.

Há ainda custo institucional. Quanto mais uma organização depende de um modelo proprietário de alta capacidade, mais processos, dados, formação e contratos são adaptados ao ecossistema do fornecedor. A substituição torna-se difícil. O aperfeiçoamento funcional pode produzir dependência, opacidade e perda de autonomia.

5 NOVAS CATEGORIAS PARA UM PROBLEMA INTEGRADO

5.1 Computação, capacidade do modelo e capacidade algorítmica

Precisão conceitual exige separar três níveis. **Computação** designa recursos físicos e operacionais empregados para treinar e executar sistemas: processadores, memória, armazenamento, energia, redes e tempo de máquina. **Capacidade do modelo** refere-se ao desempenho funcional do artefato em determinadas tarefas, sob condições e métricas definidas. **Capacidade algorítmica**, no sentido deste artigo, é mais ampla: representa o conjunto de possibilidades cognitivas e operacionais efetivamente disponibilizadas em um contexto sociotécnico.

A capacidade algorítmica inclui modelo, computação por solicitação, janela de contexto, acesso a dados, busca, ferramentas, tempo de execução, prioridade, disponibilidade, segurança, auditabilidade, suporte, continuidade e possibilidade de modificação. Dois usuários podem interagir com o “mesmo modelo” e receber capacidades materialmente diferentes porque planos, limites, integrações e registros não são iguais.⁴

Essa distinção evita antropomorfizar a IA. “Capacidade cognitiva” não significa consciência ou inteligência humana. Designa a aptidão funcional de um sistema para mediar tarefas de conhecimento, interpretação, previsão, recomendação e decisão que afetam a cognição e a ação humanas.

5.2 Economia de aperfeiçoamento

Propõe-se empregar **economia de aperfeiçoamento** como categoria analítica e normativa do incremento marginal de capacidade algorítmica. Sua pergunta não é se determinado sistema é avançado em abstrato, mas se o próximo incremento produz benefício adicional demonstrável diante de seus custos materiais, distributivos, fiscais, concorrenciais e de dependência.⁵

O conceito não se confunde com economia da inovação, que examina incentivos e processos de mudança tecnológica, nem com IA frugal, voltada ao desenvolvimento sob restrição, nem com *right-sizing*, dirigido à adequação entre modelo e tarefa. A economia de aperfeiçoamento incorpora essas contribuições, mas acrescenta três exigências:

1. o ganho marginal deve ser demonstrado em relação à finalidade concreta, não apenas a benchmark geral;
2. custos e externalidades devem ser avaliados ao longo da cadeia e distribuídos entre beneficiários e afetados;
3. a decisão sobre o incremento deve ser justificável, revisável e contestável.

A expressão “economia” não indica redução de todos os valores a preço. Direitos fundamentais, segurança, soberania e danos ecológicos podem operar como limites, e não como parcelas monetizáveis de uma equação. O método adequado é deliberação estruturada sob restrições, não simples cálculo de custo-benefício.

⁴“Capacidade algorítmica” não é utilizada como sinônimo de inteligência, consciência ou valor humano. Trata-se de categoria sociotécnica de acesso funcional.

⁵“Economia de aperfeiçoamento” é empregada aqui em sentido próprio. Não pressupõe que aperfeiçoamentos sejam escassos apenas por razões naturais; examina como custos, instituições e poder orientam sua produção e distribuição.

5.3 Economia política da capacidade algorítmica

Define-se **economia política da capacidade algorítmica** como perspectiva de análise das relações de poder que determinam a produção, propriedade, distribuição, tarifação, limitação e destinação da capacidade computacional e funcional incorporada aos sistemas de IA, incluindo a distribuição de benefícios, riscos e externalidades ao longo da cadeia sociotécnica.

Essa categoria pergunta:

- quem controla minerais, chips, nuvem, modelos, interfaces e dados;
- quem define benchmarks, níveis de serviço, preços e prioridades;
- quais atores podem pesquisar, auditar, modificar e substituir sistemas;
- quais territórios fornecem energia, água, incentivos e trabalho;
- quem captura renda, propriedade intelectual e poder de agenda;
- quem suporta falhas, dependência, poluição, resíduos e deslocamentos sociais.

O foco não é demonizar a empresa privada. Inovação empresarial, investimento e escala podem gerar benefícios sociais relevantes. O problema surge quando poder econômico se converte em autoridade quase institucional sem deveres equivalentes de transparência, universalidade, continuidade e prestação de contas. Plataformas e serviços de nuvem podem assumir funções de ordenação antes associadas a instituições públicas (LEHDONVIRTA, 2022). Empresas passam a administrar infraestruturas cognitivas essenciais, mas suas decisões permanecem orientadas primariamente por contratos e estratégia competitiva.

5.4 Privatização do limiar de suficiência

Há **privatização do limiar de suficiência** quando agentes privados, em razão do domínio de infraestrutura, modelos ou interfaces, definem de modo unilateral e pouco contestável o nível de capacidade, qualidade, profundidade, tempo, ferramentas e auditabilidade considerado suficiente para pessoas, organizações ou serviços.

O fenômeno não exige exclusão total. Pode ocorrer por meio de planos de assinatura, limites de contexto, filas, redução silenciosa de modelo, *routing* automático, restrição de ferramentas, preços variáveis, indisponibilidade regional ou contratos que dificultam auditoria e migração. O usuário conserva acesso formal à IA, mas não controla as condições materiais desse acesso.

Privatização não significa apenas propriedade privada. Um fornecedor contratado pelo Estado pode definir, por arquitetura ou contrato, a capacidade disponível a escolas, hospitais,

tribunais e órgãos administrativos. Nessa hipótese, a decisão pública foi transferida ao desenho técnico sem debate equivalente ao que seria exigido para uma regra formal.

5.5 Suficiência algorítmica distributiva

Para impedir que a suficiência legitime tecnologia inferior para quem possui menor renda ou poder, formula-se o princípio da **suficiência algorítmica distributiva**:

Cada sistema deve mobilizar apenas a capacidade materialmente necessária à finalidade legítima, mas nenhuma pessoa ou instituição pode receber menos capacidade, qualidade, segurança ou auditabilidade do que os direitos envolvidos, o risco da decisão e a finalidade pública exigem.

O princípio combina teto e piso, mas ambos possuem estatutos diferentes. O **piso** é uma garantia forte: protege o conjunto de funções indispensáveis ao direito, ao risco e à finalidade. O **teto** é um limiar de justificação: acima dele, aperfeiçoamentos continuam possíveis, porém o proponente assume o ônus de demonstrar ganho marginal e governabilidade dos custos. Não se sustenta que toda vantagem acima do piso seja moralmente irrelevante, crítica clássica ao suficientarismo; sustenta-se que ela não possui prioridade automática sobre déficits abaixo do piso e sobre externalidades não internalizadas.

Para evitar arbitrariedade, o limiar não é fixado por número de parâmetros, marca ou classe social. Sua “moeda” é a capacidade funcional situada: desempenho na tarefa, contexto, ferramentas, segurança, auditabilidade, disponibilidade, continuidade e possibilidade de revisão. A diferença entre capacidades deve decorrer de requisitos verificáveis da tarefa e do risco, não do valor social atribuído ao usuário. Uma consulta recreativa pode ser atendida por sistema modesto, mas isso não autoriza negar informação crítica de qualidade a quem não pode pagar. Uma tarefa estatal de alto impacto pode exigir auditabilidade e validação superiores, mas não necessariamente o maior modelo disponível.

Suficiência distributiva também exige que as externalidades não sejam deslocadas sistematicamente para territórios e grupos sem poder de decisão. Um sistema só é “eficiente” de forma socialmente relevante se a economia interna do fornecedor não depender da externalização de energia, água, resíduos, trabalho precário ou risco público.

5.6 Soberania sobre o limiar

Soberania sobre o limiar é a capacidade jurídica, institucional e material de participar da definição, revisão e fiscalização dos critérios que determinam quanta capacidade algorítmica é necessária, para qual finalidade, a que custo, com quais dependências e sob quais salvaguardas.

Ela possui quatro dimensões:

- **normativa:** poder de definir critérios, padrões e direitos;
- **material:** acesso a computação, energia, dados, redes e equipes;
- **tecnológica:** capacidade de auditar, modificar, interoperar e substituir;
- **decisória:** liberdade real para escolher ou abandonar ecossistemas sem paralisar serviços essenciais.

Soberania não requer autarquia tecnológica. Nenhum país controla isoladamente toda a cadeia de semicondutores e IA. O requisito é reduzir pontos de indispensabilidade, diversificar com interoperabilidade, preservar conhecimento interno e manter capacidade de saída. Diversificar fornecedores sem construir substituíbilidade pode apenas diversificar dependências.

5.7 Estatuto e alcance das categorias propostas

As cinco expressões não são apresentadas como descobertas de objetos inexistentes na literatura. Cada uma possui antecedentes: incremento marginal e eficiência; economia política da infraestrutura; poder privado e *private ordering*; suficientarismo; governança da computação; e soberania digital. A contribuição reivindicada é mais delimitada e verificável.

Primeiro, a unidade de análise deixa de ser apenas o chip, o modelo ou o consumo e passa a ser a **capacidade funcional efetivamente entregue**. Segundo, o incremento é submetido simultaneamente a ciclo de vida, distribuição, direitos e dependência. Terceiro, o limiar é tratado como decisão institucional contestável: é preciso identificar quem o fixa, com qual evidência, sob qual ônus e com qual possibilidade de recurso. Quarto, soberania inclui a capacidade de abandonar um ecossistema sem paralisar a função atendida.

Essa delimitação produz hipóteses refutáveis. A “privatização do limiar” perde utilidade se decisões de capacidade forem transparentes, multilateralmente governadas, recorríveis e substituíveis. A “economia de aperfeiçoamento” falha se não distinguir incremento justificável de mera preferência estética. A “suficiência algorítmica distributiva” torna-se opressiva se classificar pessoas em vez de requisitos funcionais. E a “soberania sobre o limiar” vira retórica se não puder ser medida por auditabilidade, conhecimento interno, interoperabilidade e tempo de saída.

6 CAPACIDADE ALGORÍTMICA E INFLUÊNCIA COGNITIVA ESTRUTURAL

6.1 A decisão começa antes do resultado

Em *Autonomia decisória na era da IA*, a **influência cognitiva estrutural** é definida como a atuação das arquiteturas informacionais sobre as condições de percepção, compreensão, formulação do problema e escolha antes que a decisão final seja exteriorizada (GENOVA, 2026). A categoria dialoga com a arquitetura de escolha (THALER; SUNSTEIN, 2008), com a regulação por *hypernudges* adaptativos (YEUNG, 2017) e com estudos sobre complacência e viés de automação (PARASURAMAN; MANZEY, 2010), mas não se reduz a eles. Seu objeto é a organização duradoura das condições do decidir, inclusive dependência material e perda de alternativas institucionais. Ela não exige comando direto, engano deliberado ou coerção. Opera pela organização do campo: o que aparece, em que ordem, com qual linguagem, quais alternativas são sugeridas e quais permanecem invisíveis.

A distribuição de capacidade algorítmica acrescenta uma camada material a essa tese. Não basta perguntar como um modelo formula a resposta. É necessário perguntar qual modelo foi disponibilizado, quanto contexto recebeu, que bases consultou, que ferramentas pôde utilizar, quanto tempo teve para processar e se o usuário poderia exigir outra rota. A arquitetura de alocação antecede o conteúdo.

Quando um provedor encaminha uma solicitação a um sistema reduzido, limita buscas, encurta contexto ou desativa ferramentas, ele não apenas economiza computação. Delimita o universo de respostas possíveis. Em atividades jurídicas, médicas, científicas ou administrativas, essa diferença pode alterar hipóteses consideradas, exceções percebidas e evidências recuperadas.

6.2 A capacidade como condição do pensável institucional

Organizações pensam por meio de rotinas, documentos, indicadores, profissionais e infraestruturas. Ao incorporar IA, passam a depender da capacidade que o sistema torna disponível. Se um município, hospital ou tribunal utiliza modelo que não pode auditar, modificar ou substituir, parte de seu campo cognitivo institucional fica condicionada por decisões externas.

Esse condicionamento não se reduz ao erro de uma resposta. Pode afetar:

- quais problemas são formalizados como tratáveis;

⁶A expressão foi desenvolvida pelo autor para examinar condições estruturais anteriores à decisão manifesta. Não se confunde com viés de automação, persuasão episódica ou manipulação individual, embora possa interagir com esses fenômenos.

- quais evidências são consideradas recuperáveis;
- quais categorias e classificações se tornam padrão;
- quanto tempo é reservado à revisão humana;
- quais alternativas parecem economicamente viáveis;
- quais capacidades profissionais deixam de ser exercidas internamente.

O efeito é cumulativo. Equipes adaptam sua linguagem ao sistema, bases são reestruturadas para a plataforma, contratos renovam dependências e profissionais perdem familiaridade com processos anteriores. A influência cognitiva estrutural passa a operar também pela assimetria de saída: a organização formalmente decide, mas já não consegue decidir nas mesmas condições sem o fornecedor.

6.3 Estratificação por interface

Mercados digitais já distribuem velocidade, armazenamento, visibilidade e suporte por planos. Com IA, a segmentação alcança uma capacidade que participa da produção de conhecimento. Versões gratuitas, profissionais, empresariais e governamentais podem oferecer diferentes modelos, janelas de contexto, ferramentas, garantias e prioridade. Diferenças de preço são normais em mercados; tornam-se problema público quando a infraestrutura funciona como condição de exercício de direitos ou prestação de serviços essenciais.

A estratificação pode ser invisível. Duas interfaces semelhantes ocultam modelos diferentes; mudanças são realizadas sem registro compreensível; o usuário não recebe justificativa sobre o sistema que tratou sua solicitação. A aparência de igualdade de acesso convive com desigualdade material de capacidade.

Por isso, transparência não deve limitar-se a informar que “há IA”. Deve alcançar, na medida compatível com segurança e segredo legítimo, a classe de capacidade empregada, restrições relevantes, critérios de roteamento, alterações materiais e consequências previsíveis. Em sistemas de alto impacto, registros precisam permitir reconstruir qual configuração participou da decisão.

7 A FRONTEIRA ANTIELITISTA: LIMITAR INCREMENTOS, NÃO PESSOAS

7.1 O risco moral de classificar usos e usuários

O exemplo intuitivo que contrapõe uma consulta recreativa a uma pesquisa médica parece demonstrar que a segunda mereceria um sistema superior. Mas, convertido em política

geral, esse raciocínio cria uma autoridade encarregada de julgar finalidades, usuários e valores. Quem define o que é trivial? Uma pergunta sobre astrologia pode integrar pesquisa cultural, investigação de fraude, cuidado psicológico ou simples lazer legítimo. O Estado e a plataforma não devem exigir demonstração de produtividade para reconhecer dignidade digital.

Além disso, categorias aparentemente neutras reproduzem poder. Instituições bem financiadas descrevem seus projetos com linguagem compatível com editais, métricas e prioridades oficiais. Comunidades, pesquisadores independentes e problemas locais podem parecer menos relevantes porque não controlam os códigos de legitimidade. Uma governança baseada em “mérito do usuário” tenderia a premiar posições já privilegiadas.

7.2 O objeto correto da avaliação

A unidade de avaliação deve ser o **incremento de capacidade em relação a uma finalidade**, e não o valor da pessoa. O procedimento pergunta:

- qual resultado mínimo é necessário;
- que riscos decorrem de desempenho insuficiente;
- qual ganho adicional a capacidade superior oferece;
- quais recursos e externalidades esse ganho exige;
- existe alternativa menos intensiva funcionalmente equivalente;
- a redução de capacidade afetará direitos, segurança ou possibilidade de contestação.

A pergunta recreativa não precisa ser proibida nem degradada por identidade do usuário. Pode ser processada por arquitetura eficiente porque seus requisitos técnicos são modestos e porque um sistema maior não demonstrou benefício relevante. Se a mesma interação envolver risco, pesquisa ou necessidade específica, o limiar muda. O critério é contextual, não moralizante.

7.3 Piso universal e diferenciação justificável

Uma governança não elitista exige um piso universal composto por:

1. segurança compatível com o contexto;
2. informação clara sobre limitações relevantes;
3. proteção de dados e confidencialidade;
4. ausência de discriminação ilícita;
5. possibilidade de correção e contestação;

6. acesso a capacidade adequada quando direitos ou serviços essenciais estiverem envolvidos;
7. alternativas acessíveis para pessoas com deficiência e grupos vulnerabilizados;
8. preservação de canais humanos quando a decisão não puder ser integralmente automatizada.

Acima do piso, diferenças podem ser justificadas por complexidade, risco, segurança, privacidade, urgência e finalidade pública. Elas não podem decorrer secretamente de renda, localização ou perfil comportamental quando produzirem impacto relevante. A oferta comercial de níveis premium permanece possível, mas não deve definir o padrão material de direitos.

7.4 Alto impacto não significa modelo máximo

Uma confusão frequente associa importância da decisão ao tamanho do modelo. Sistemas de alto impacto exigem confiabilidade, validação, rastreabilidade, segurança e governança. Um modelo geral de fronteira pode ser inadequado por opacidade, instabilidade ou dificuldade de reprodução. Em certos contextos, regra verificável, modelo especializado ou processo humano estruturado oferece melhor proteção.

O limiar deve ser suficiente para a finalidade e o risco, não maximalista. A maior capacidade comercial disponível não é sinônimo de maior adequação jurídica.

8 O TESTE DO LIMIAR DE SUFICIÊNCIA ALGORÍTMICA

8.1 Estatuto, campo de aplicação e fórmula

Propõe-se o **Teste do Limiar de Suficiência Algorítmica (TLSA)** como estrutura heurística analítico-normativa para organizar a justificação de escolhas sociotécnicas. A denominação “teste” tem função mnemônica e procedimental: nesta etapa da pesquisa, não designa instrumento psicométrico, métrica validada nem algoritmo capaz de gerar sozinho uma resposta correta. Sua consistência interavaliadores, seus custos de aplicação e sua sensibilidade a diferentes setores ainda precisam ser submetidos a estudos empíricos.

O TLSA incide sobre três classes de decisão:

- **desenvolvimento:** se deve ser produzido ou escalado novo nível de capacidade;
- **provisionamento:** como a capacidade será distribuída entre instituições, territórios e serviços;
- **execução:** qual combinação de modelo, ferramentas, dados e supervisão deve tratar determinada tarefa.

Sua fórmula é:

Nem mais capacidade do que a finalidade justifica; nem menos capacidade do que o risco e os direitos exigem.

O TLSA não converte dignidade, água, carbono, soberania, custo e acurácia em moeda única. Algumas exigências admitem ponderação; outras operam como pisos, vedações ou condições de validade. Ganho financeiro, conveniência ou melhoria média de desempenho não compensam automaticamente violação de direito fundamental, discriminação ilícita, dano ambiental juridicamente intolerável ou supressão da possibilidade efetiva de revisão.

8.2 Relação com a proporcionalidade

O TLSA não pretende renomear a proporcionalidade. Ele a toma como matriz jurídica e a especializa para decisões sobre capacidade algorítmica. Na reconstrução usual, a proporcionalidade examina a legitimidade da finalidade, a adequação do meio, a necessidade diante de alternativas menos gravosas e a proporcionalidade em sentido estrito entre benefícios e restrições (ALEXY, 2003; BARAK, 2012; ÁVILA, 2021). O TLSA conserva essa gramática e acrescenta quatro focos que a aplicação abstrata pode deixar subespecificados:

1. o **ciclo de vida material**, incluindo computação, energia, água, equipamentos e efeitos territoriais;
2. o **piso distributivo**, que impede que eficiência se converta em capacidade insuficiente para direitos e serviços essenciais;
3. a **dependência sociotécnica**, abrangendo interoperabilidade, conhecimento interno, continuidade e saída;
4. a **revisão dinâmica**, porque modelos, custos, riscos e alternativas mudam durante a vigência da decisão.

A finalidade legítima corresponde à definição do problema; a adequação exige desempenho demonstrado na tarefa; a necessidade impõe comparação com configurações menos intensivas; e a ponderação examina benefício marginal, externalidades, distribuição e dependência. O componente suficiente acrescenta uma direção dupla. A proporcionalidade normalmente contém o excesso estatal; aqui ela também precisa identificar **insuficiência estrutural** quando economia, segmentação comercial ou escassez administrada reduzem proteção abaixo do piso devido.

Essa relação limita a pretensão de originalidade: o TLSA é uma **especialização sociotécnica e distributiva da proporcionalidade**, informada pela avaliação de impacto,

pela contratação por ciclo de vida e pela governança de risco. Seu valor será demonstrado apenas se tornar visíveis evidências e conflitos que os procedimentos existentes não capturam de modo estável.

8.3 As dez perguntas do teste

1. Finalidade legítima e problema definido

Qual problema concreto será enfrentado? A finalidade é lícita, específica e compatível com a missão da instituição? Sistemas não devem ser adquiridos ou ampliados apenas porque a capacidade existe. Formulações vagas — “modernizar”, “usar IA”, “não ficar para trás” — não constituem motivação suficiente.

2. Desempenho mínimo necessário

Quais requisitos de qualidade, segurança, latência, disponibilidade, explicabilidade e auditabilidade são indispensáveis? O mínimo deve ser definido antes da comparação com fornecedores, para reduzir a possibilidade de que a oferta disponível fabrique a necessidade.

3. Benefício marginal demonstrável

Qual ganho verificável resulta do nível adicional de capacidade? O benefício aparece em tarefa real, população relevante e condições de operação, ou apenas em *benchmark* geral? Diferenças pequenas podem ser decisivas quando direitos estão em jogo, mas não devem ser presumidas socialmente valiosas sem demonstração contextual.

4. Alternativa menos intensiva

Existe modelo menor, sistema especializado, processamento local, recuperação de informação, regra determinística, melhoria de dados ou reorganização do trabalho capaz de atingir o piso requerido? A comparação deve considerar qualidade total e ciclo de vida, não somente parâmetros ou preço por consulta.

5. Impacto material de ciclo de vida

Quais são os efeitos sobre energia, água, emissões, minerais, equipamentos, rede e resíduos? Quem controla os dados e qual incerteza acompanha a estimativa? A ausência de informação deve ser registrada como risco de governança, não como impacto zero.

6. Distribuição de benefícios e ônus

Quem recebe produtividade, renda, conhecimento e qualidade de serviço? Quem suporta tarifas, escassez local, trabalho precário, vigilância, falhas e descarte? A análise deve incluir territórios e grupos afetados, não apenas contratante e fornecedor.

7. Dependência, interoperabilidade e saída

Dados, registros, modelos auxiliares e processos podem ser migrados? Existem padrões abertos, documentação e pessoal interno? Qual é o tempo e o custo de abandonar o ecossistema sem interromper a função? A capacidade de saída integra a própria adequação do objeto.

8. Transparência, auditoria e contestação

É possível reconstruir qual classe de sistema foi utilizada, quais limitações importaram e como o resultado pode ser revisto? Sigilos legítimos podem graduar a publicidade, mas não eliminar a evidência necessária à fiscalização e à defesa.

9. Piso de direitos e não discriminação

A economia de recursos reduzirá proteção, acesso ou qualidade abaixo do exigido por direitos, risco e finalidade pública? Os critérios de alocação produzem desvantagem injustificada por renda, território, raça, gênero, deficiência, posição institucional ou perfil?

10. Reversibilidade e revisão periódica

A decisão pode ser suspensa, reduzida, ampliada ou substituída? Quais indicadores acionam revisão e quem decide? Como modelos, custos e impactos mudam rapidamente, suficiência é limiar revisável, não número permanente.

8.4 Competência, ônus da prova e padrão de evidência

A instância que autoriza o desenvolvimento, a contratação ou o uso conserva a responsabilidade de decidir. No setor público, isso significa que a autoridade competente não pode transferir ao fornecedor a definição da finalidade, a avaliação do risco ou a conclusão sobre suficiência. Uma equipe interdisciplinar instrui o processo; controle interno, encarregado de dados, unidade ambiental ou autoridade setorial participam conforme a matéria; e revisão independente torna-se necessária nos regimes reforçados.

O **ônus inicial de demonstração** cabe a quem propõe o incremento ou controla a informação pertinente. O fornecedor deve comprovar alegações sobre desempenho, consumo, segurança, mudanças de versão e portabilidade; a unidade demandante deve justificar finalidade, requisito funcional e população afetada; a instituição deve documentar comparação, decisão e supervisão. Não se exige certeza impossível. O padrão é o de **evidência adequada ao risco, à escala e à reversibilidade**, com hipóteses, intervalos e lacunas expressamente declarados.

Quando informação indispensável permanece exclusivamente sob controle do proponente, sua ausência não pode beneficiá-lo. Em baixo risco, ela pode conduzir a piloto limitado ou cláusula de produção posterior. Em alto impacto, pode impedir a contratação ou o uso. O sigilo empresarial justifica acesso restrito, compromisso de confidencialidade, resumo público e auditoria por terceiro; não autoriza uma “caixa-preta probatória” na qual o poder público assuma o risco sem conhecer o objeto.

8.5 Regra de decisão, vedações e remédios

O nível de capacidade é justificável quando:

1. atende ao piso de direitos, segurança e desempenho;
2. apresenta ganho marginal relevante para a finalidade;
3. não existe alternativa menos intensiva funcionalmente equivalente;
4. impactos e dependências são proporcionais, monitoráveis e governáveis;
5. benefícios e ônus não são distribuídos de forma injustificada;
6. a decisão permanece transparente na medida juridicamente exigível, contestável e reversível.

O resultado do TLSA não é apenas “aprovar” ou “rejeitar”. A autoridade pode: **aprovar; aprovar sob condições; autorizar piloto limitado; exigir informação ou nova comparação; reduzir ou ampliar capacidade; suspender; ou recusar.** As condições podem envolver restrição de finalidade, revisão humana, auditoria, limite temporal, arquitetura modular, métricas, compensação territorial e plano de saída.

Constituem impedimentos mínimos: insuficiência abaixo do piso de direitos; discriminação ilícita não mitigável; impossibilidade de revisão em função de alto impacto; ausência de continuidade e saída em serviço essencial; uso vedado pela legislação; e dano grave não mitigável ou juridicamente intolerável. Nesses casos, melhoria média de desempenho não resolve o defeito.

A decisão deve identificar responsável, evidências, divergências, condições, indicadores e data de revisão. Pessoas afetadas precisam de canal de contestação compatível com o impacto. Mudança material de modelo, fornecedor, finalidade, dados, região de processamento ou política de roteamento reabre o exame. Os remédios incluem correção, reproprocessamento, revisão humana, comunicação aos afetados, reparação, substituição e descontinuação.

8.6 Matriz operacional

Quadro 3 — Matriz operacional do Teste do Limiar de Suficiência Algorítmica

Dimensão	Evidência mínima	Sinal de alerta	Resposta possível
Finalidade	problema, público e resultado definidos	aquisição orientada por moda ou fornecedor	reformular ou não contratar
Desempenho e margem	teste em tarefa e população relevantes	vantagem apenas em <i>benchmark</i> genérico	piloto comparativo ou recusa da alegação
Alternativa	comparação com opções menores e não algorítmicas	modelo de fronteira como padrão automático	exigir arquitetura proporcional
Ciclo de vida	energia, água, hardware, origem e incerteza	dados ausentes ou excessivamente agregados	mensuração, auditoria, condicionante ou recusa
Distribuição	mapa de beneficiários, afetados e território	ônus local e benefício externalizado	redesenho, compensação ou não adoção
Direitos	avaliação de impacto e piso de proteção	redução de qualidade para grupo vulnerabilizado	vedar redução, corrigir e garantir revisão humana
Dependência	portabilidade, interoperabilidade e plano de saída testado	bloqueio contratual ou técnico	modularidade, depósito técnico e padrões abertos
Governança	registro, auditoria, recurso e revisão	roteamento ou mudança silenciosa	notificação, trilha de decisão e sanção

Fonte: elaboração do autor.

8.7 Intensidade proporcional

Para evitar burocratização, o TLISA deve possuir três regimes:

Quadro 4 — Regimes de intensidade do TLISA

Regime	Situação típica	Exigência
Simplificado	uso rotineiro, reversível e de baixo impacto	finalidade, classe de capacidade, proteção de dados e revisão amostral
Completo	contratação coletiva, integração organizacional ou impacto relevante	dez perguntas, comparação de alternativas, avaliação de impacto e plano de saída
Reforçado	infraestrutura de grande escala, modelo de fronteira, serviço essencial ou risco sistêmico	revisão independente, participação dos afetados, dados de ciclo de vida, monitoramento contínuo e autorização por instância competente

Fonte: elaboração do autor.

A classificação considera escala, duração, irreversibilidade, direitos, externalidades e dependência. Volume elevado de operações simples pode justificar regime superior por efeito acumulado. Inversamente, uma tarefa tecnicamente complexa, mas experimental, isolada e reversível, pode ser controlada sem procedimento público extenso. O enquadramento e eventual reclassificação devem ser motivados.

8.8 Aplicação em compras públicas

Contratações públicas de IA oferecem campo imediato ao TLISA. A Lei nº 14.133/2021 vincula a licitação ao resultado mais vantajoso, inclusive quanto ao ciclo de vida; prevê desenvolvimento nacional sustentável; exige planejamento; e atribui à alta administração a governança das contratações, com gestão de riscos e controles internos (arts. 5º, 11, 18 e 169) (BRASIL, 2021). Esses comandos permitem superar a comparação estreita por preço ou notoriedade do fornecedor.

Um edital orientado pelo TLISA deve especificar:

- requisitos mínimos e cenários de teste;
- classes de capacidade e critérios de roteamento;
- consumo e impactos reportáveis;
- registros necessários à auditoria;
- portabilidade de dados, configurações e conhecimento;
- níveis de serviço e continuidade;
- procedimento de contestação;
- inspeção independente sob regime de confidencialidade;
- revisão diante de mudança material;
- plano testável de saída e substituição.

Contratar “acesso à IA” sem essas condições equivale a permitir que o fornecedor redefina o objeto durante a execução.

8.9 Caso-tipo: assistência generativa no Poder Judiciário

Considere-se um **caso-tipo hipotético**, construído para testar o procedimento e não para descrever contratação real: um tribunal pretende contratar assistência generativa para pesquisa em acervo autorizado, resumo de processos e preparação de minutas. A comparação preliminar contém três arquiteturas:

- **A — modelo geral de fronteira em nuvem externa**, com ampla capacidade e atualização contínua;
- **B — modelo de porte intermediário com recuperação de informação em base institucional, execução segregada e possibilidade de substituição modular**;
- **C — busca semântica, extração e modelos determinísticos de documentos, sem geração livre para determinadas tarefas**.

A Resolução CNJ nº 615/2025 exige governança, avaliação preliminar, classificação de risco, monitoramento, transparência, supervisão humana e avaliação de impacto para

soluções de alto risco. Ela veda, entre outros usos, soluções que não permitam revisão humana ou gerem dependência absoluta do resultado (arts. 1º a 14) (CONSELHO NACIONAL DE JUSTIÇA, 2025). O TLSA não substitui essas obrigações; organiza a pergunta adicional sobre qual capacidade é necessária e quem controla sua alteração.

Quadro 5 — Aplicação ilustrativa do TLSA ao caso-tipo

Questão	Evidência a produzir	Consequência provisória
Finalidade	separar pesquisa, resumo, minuta e decisão; definir o que permanece vedado à automação	impedir contratação genérica para “apoio jurisdicional”
Piso funcional	testes com classes processuais, linguagem e documentos representativos; taxa de erro crítico; revisão	arquitetura distinta pode ser exigida por tarefa
Benefício marginal	comparação cega entre A, B e C em cenário real, com custo total e tempo de revisão	A não vence apenas por desempenho médio superior
Alternativa	verificar se C satisfaz extração e busca e se B satisfaz resumo com fontes	roteamento pode reservar A apenas a hipóteses justificadas
Materialidade	localização, consumo, retenção, equipamentos e volume estimado	condicionar escala e exigir relatório verificável
Direitos	sigilo, dados pessoais, vieses, contraditório, motivação e supervisão	nenhum texto gerado possui autoridade decisória própria
Dependência	exportação de dados, registros, <i>prompts</i> , índices e configuração; prazo de migração	ausência de saída impede uso essencial continuado
Governança	log, identificação de fontes, versão, incidentes, auditoria e recurso	mudança silenciosa reabre avaliação

Fonte: elaboração do autor.

Sem resultados empíricos, não é possível declarar antecipadamente a arquitetura vencedora. O exercício mostra, porém, como o TLSA altera a instrução. A arquitetura C pode ser suficiente para busca e extração; B pode ser autorizada em piloto para resumos e minutas não vinculantes, com recuperação de fontes e revisão; A pode permanecer disponível para tarefas excepcionais se o ganho marginal for demonstrado. Para matéria decisória de alto impacto, o piso sobe: rastreabilidade, validação, supervisão efetiva e contestação tornam-se condições, e determinada função pode permanecer integralmente humana.

O exemplo também contém o risco inverso. Se B ou C apresentarem erro crítico superior, cobertura linguística inadequada ou barreira de acessibilidade, a economia de recursos não justifica seu uso. O TLSA pode mandar **ampliar** capacidade. Essa simetria impede que a suficiência seja convertida em austeridade tecnológica para quem depende do serviço público.

9 FUNDAMENTOS JURÍDICOS DE UMA GOVERNANÇA DO LIMIAR

9.1 Estatuto jurídico da proposta

O ordenamento brasileiro não contém, em julho de 2026, um dever geral denominado “suficiência algorítmica” nem autoriza, por essa expressão, racionamento universal de capacidade computacional. A validade jurídica da proposta depende de separar três planos:

1. **de *lege lata*:** deveres já extraíveis de normas vigentes, como finalidade, necessidade, planejamento, prevenção, transparência, não discriminação, governança de risco, ciclo de vida e responsabilização;
2. **integração funcional:** emprego do TLISA como roteiro probatório ou contratual para cumprir deveres existentes, sem ampliar por analogia o âmbito de incidência de uma lei;
3. **de *lege ferenda*:** obrigações novas sobre impacto material, alocação de capacidade, interoperabilidade, computação pública e governança do limiar, dependentes de lei, regulação setorial ou cláusula contratual válida.

Essa distinção evita dois excessos: apresentar o TLISA como direito positivo já consolidado ou reduzir o ordenamento vigente à completa irrelevância. Há bases jurídicas reais, mas sua combinação e extensão precisam respeitar competência, legalidade e devido processo.

9.2 Constituição: dignidade, igualdade, desenvolvimento e soberania

A Constituição oferece fundamento transversal. A dignidade da pessoa humana (art. 1º, III), os objetivos de redução das desigualdades (art. 3º, III), a igualdade, a intimidade e a proteção de dados pessoais (art. 5º, *caput*, X e LXXIX) impedem que funções essenciais sejam submetidas a segmentação opaca e discriminatória. O reconhecimento constitucional da proteção de dados pela Emenda Constitucional nº 115/2022 reforça que infraestrutura informacional e autonomia não são questões exclusivamente comerciais (BRASIL, 1988; BRASIL, 2022).

A ordem econômica combina valorização do trabalho e livre iniciativa com função social da propriedade, livre concorrência, defesa do consumidor, defesa do meio ambiente e redução das desigualdades (arts. 170, III a VII). O Estado atua como agente normativo e regulador da atividade econômica (art. 174), enquanto ciência, tecnologia e inovação devem orientar-se, entre outros fins, à solução dos problemas brasileiros e ao desenvolvimento do sistema produtivo nacional (arts. 218 e 219).

Esses dispositivos não criam direito individual ao maior modelo disponível, tampouco proíbem abstratamente sistemas intensivos. Sustentam, de *lege lata*, decisões públicas motivadas por direitos, sustentabilidade, concorrência e desenvolvimento. A criação de cotas gerais de computação ou obrigações privadas de acesso, por outro lado, exigirá base legal específica e desenho compatível com propriedade, livre iniciativa, igualdade e competências federativas.

9.3 Direito ambiental: prevenção, licenciamento e ciclo de vida

O art. 225 da Constituição reconhece o direito ao meio ambiente ecologicamente equilibrado e impõe ao poder público e à coletividade o dever de defendê-lo e preservá-lo. A Lei nº 6.938/1981 estabelece a Política Nacional do Meio Ambiente, seus instrumentos, o licenciamento de atividades efetiva ou potencialmente poluidoras e a responsabilidade civil objetiva por dano ambiental (arts. 2º, 4º, 9º, IV, 10 e 14, § 1º) (BRASIL, 1981).

Prevenção incide quando impactos são identificáveis; precaução orienta a decisão diante de incerteza relevante e possibilidade de dano grave. Em centros de dados e infraestrutura de IA, a consequência não é licenciar “o modelo” como entidade abstrata. É avaliar empreendimentos, captação de água, energia, equipamentos, emissões, resíduos, localização e efeitos cumulativos segundo a legislação aplicável e as competências dos órgãos envolvidos.

O princípio do poluidor-pagador exige internalização, mas não transforma pagamento em licença para degradar. Compensação pode ser inadequada diante de limite territorial, dano irreversível ou distribuição incompatível com direitos. O TLSA funciona aqui como integração probatória: pergunta se o incremento que produz a externalidade foi demonstrado e se existe alternativa funcional menos onerosa. Novos deveres nacionais padronizados de divulgação de energia e água para toda a cadeia de IA permanecem matéria de *lege ferenda*.

9.4 Proteção de dados: incidência direta e analogia limitada

O art. 6º, III, da Lei nº 13.709/2018 define necessidade como limitação do tratamento ao mínimo necessário à finalidade, com dados pertinentes, proporcionais e não excessivos. Finalidade, adequação, transparência, segurança, prevenção, não discriminação e responsabilização também vinculam operações sujeitas à LGPD. Decisões unicamente automatizadas que afetem interesses do titular atraem o direito de revisão do art. 20, nos limites de sua redação e interpretação (BRASIL, 2018).

Quando roteamento, treinamento, registro ou inferência tratam dados pessoais, esses comandos incidem diretamente. O Supremo Tribunal Federal, ao referendar as medidas cautelares nas ADIs nº 6.387, 6.388, 6.389, 6.390 e 6.393, suspendeu a MP nº 954/2020 e exigiu justificação compatível com finalidade, necessidade e proteção adequada em compartilhamento massivo de dados. O precedente não tratou de capacidade algorítmica, mas confirma que conveniência administrativa e finalidade genérica não bastam diante de riscos informacionais (SUPREMO TRIBUNAL FEDERAL, 2020).

Minimização de dados, contudo, não equivale a minimização de computação. Um sistema pode tratar poucos dados com grande custo material ou exigir processamento adicional para proteger segurança e acessibilidade. A LGPD não distribui chips, energia ou modelos. Sua estrutura é analogia metodológica para o limiar, não fundamento para ampliar, sem lei, seu objeto.⁷

9.5 Contratação pública: planejamento, ciclo de vida e gestão de riscos

A Lei nº 14.133/2021 oferece a base positiva mais imediata para uso institucional do TLISA. Seus princípios incluem planejamento, transparência, eficiência, interesse público, desenvolvimento nacional sustentável, motivação e economicidade (art. 5º). O processo licitatório deve assegurar proposta apta a gerar resultado mais vantajoso, inclusive quanto ao ciclo de vida do objeto, e evitar sobrepreço, inexecuibilidade e contratações que não atendam à necessidade (art. 11). A fase preparatória é caracterizada pelo planejamento e deve compatibilizar considerações técnicas, mercadológicas e de gestão (art. 18). A alta administração responde pela governança das contratações, e as linhas de defesa devem exercer gestão de riscos e controle preventivo (arts. 11, parágrafo único, e 169) (BRASIL, 2021).

Esses dispositivos autorizam exigir comparação funcional, custo total, segurança, portabilidade, continuidade, auditoria e saída quando pertinentes ao objeto. Não autorizam requisito sem relação com a contratação nem favorecimento disfarçado. Cada exigência precisa ser necessária, proporcional, tecnicamente motivada e aberta à competição possível.

O Acórdão nº 157/2024-Plenário do Tribunal de Contas da União, em acompanhamento de contratação de computação em nuvem, examinou medidas de mitigação de riscos adotadas por CNJ, CGU e TCU. Não validou o TLISA, mas ilustra que dependência,

⁷O princípio da necessidade da LGPD refere-se ao tratamento de dados pessoais. Sua utilização neste artigo como modelo de raciocínio proporcional não amplia, por interpretação, o objeto da lei.

segurança, governança e desenho contratual de nuvem já integram o controle de contratações de infraestrutura digital (TRIBUNAL DE CONTAS DA UNIÃO, 2024).

9.6 Concorrência, consumidor e assimetria de controle

A Lei nº 12.529/2011 protege a coletividade e reprime atos capazes de limitar a livre concorrência, dominar mercado, aumentar arbitrariamente lucros ou exercer abusivamente posição dominante (arts. 1º e 36). Integrações entre chips, nuvem, modelos, ferramentas e canais podem produzir fechamento, discriminação de acesso, autopreferência e custos de troca. A análise exige definição de mercado, poder e efeitos; concentração, por si só, não é ilícita, e eficiência pode afastar determinada caracterização (BRASIL, 2011).

A doutrina de infraestrutura essencial é controvertida e não deve ser convertida em atalho. Interoperabilidade, portabilidade, separação funcional ou acesso não discriminatório somente podem ser impostos por autoridade competente, com base normativa e instrução econômica. O TLSA pode organizar evidência de dependência, mas não substitui o teste concorrencial.

No consumo, os arts. 6º, III, 8º, 14, 20 e 31 da Lei nº 8.078/1990 protegem informação adequada, segurança, qualidade e correspondência entre oferta e entrega. Alteração silenciosa de versão, capacidade ou roteamento pode ser relevante quando muda característica prometida ou adequação do serviço. O Código de Defesa do Consumidor não garante o modelo mais caro; exige que restrições e mudanças materiais não tornem enganosa a oferta nem transfiram ao consumidor risco indevido (BRASIL, 1990).

9.7 Regulação setorial: o exemplo do Poder Judiciário

A Resolução CNJ nº 615/2025 constitui exemplo setorial concreto. O ato disciplina desenvolvimento, governança, auditoria, monitoramento e uso responsável de IA no Judiciário; adota proporcionalidade ao impacto, ciclo de vida, avaliação preliminar, classificação de risco, avaliação de impacto algorítmico, interoperabilidade, transparência, contestabilidade e supervisão humana. Os arts. 10 e 11 contêm vedações e regimes de risco; os arts. 12 a 14 estabelecem medidas de governança; e os arts. 33 e 34 preservam caráter consultivo e supervisão humana dos produtos gerados (CONSELHO NACIONAL DE JUSTIÇA, 2025).

O TLSA pode ser incorporado como roteiro complementar em estudos técnicos, avaliações e contratações do Judiciário, desde que não contrarie a Resolução nem pretenda substituir as competências do CNJ e dos tribunais. Sua função específica é comparar

capacidade, materialidade e dependência — dimensões que dialogam com risco e eficiência, mas não se confundem com elas.

9.8 Normas internacionais e o Regulamento Europeu de IA

A Recomendação da UNESCO sobre a Ética da Inteligência Artificial associa proporcionalidade e inocuidade à sustentabilidade, auditabilidade, avaliação de impacto, supervisão humana e proteção de ecossistemas (UNESCO, 2021). Por não ser tratado autoexecutável que crie, sozinho, obrigações privadas internas, funciona como orientação interpretativa e compromisso político.

O Regulamento (UE) 2024/1689 oferece referências mais precisas. O art. 51 presume capacidade de alto impacto em modelo de propósito geral quando a computação cumulativa de treinamento supera 10^{25} operações de ponto flutuante, sem reduzir a classificação a esse único indicador. O Anexo XI exige informação sobre recursos computacionais, tempo de treinamento e consumo de energia conhecido ou estimado; o art. 95, § 2º, “b”, prevê códigos voluntários para avaliar e minimizar impactos sobre sustentabilidade ambiental, inclusive técnicas eficientes de projeto, treinamento e uso (UNIÃO EUROPEIA, 2024).

O regime europeu torna a escala material juridicamente visível, mas não institui a suficiência algorítmica distributiva aqui defendida. O TLSA acrescenta, como proposta, a justificação do incremento, o piso de capacidade e a autoridade revisável sobre o limiar.

9.9 Síntese: o que já pode ser exigido e o que depende de inovação normativa

De *lege lata*, órgãos públicos podem e, conforme o caso, devem justificar finalidade, adequação, necessidade, risco, ciclo de vida, segurança, dados, concorrência, impacto ambiental e continuidade. Podem exigir evidências e cláusulas relacionadas ao objeto e aplicar vedações setoriais existentes. Empresas respondem pelos regimes jurídicos que efetivamente incidam sobre sua posição e atividade.

Dependem de desenvolvimento legislativo ou regulatório: padrões gerais de mensuração material da IA; dever universal de revelar roteamento e classes de capacidade; direito geral de escalonamento; obrigações de interoperabilidade fora de relações e setores já regulados; mecanismos compulsórios de acesso a computação; e instância nacional encarregada de governar limiares. O TLSA, portanto, é imediatamente utilizável como **método de instrução e governança voluntária ou contratual**, mas somente adquire força coercitiva na extensão da norma competente que o incorporar.

10 GOVERNANÇA INSTITUCIONAL DA CAPACIDADE

10.1 Da transparência ambiental à transparência de alocação

Relatórios de energia e emissões são necessários, mas insuficientes. A governança precisa revelar como capacidade é alocada. Em contextos relevantes, fornecedores devem documentar:

- classes de modelos e configurações disponíveis;
- critérios que encaminham solicitações;
- limitações de contexto, ferramentas e tempo;
- alterações materiais de versão;
- métricas de qualidade por tarefa;
- consumo estimado e metodologia;
- incidentes e degradações;
- possibilidade de revisão por sistema superior ou por pessoa.

A divulgação pode ser graduada para proteger segurança e propriedade intelectual. O mínimo é permitir que auditor, regulador e instituição contratante reconstruam decisões e comparem entrega com obrigação.

10.2 Governança independente do limiar

O limiar não deve ser definido exclusivamente pelo fornecedor nem por uma autoridade central sem controle. Recomenda-se governança policêntrica, inspirada na ideia de que recursos comuns complexos exigem regras em múltiplos níveis, participação e monitoramento (OSTROM, 1990).

Uma estrutura institucional pode reunir:

1. órgãos ambientais e energéticos, para impactos territoriais;
2. autoridades de proteção de dados e concorrência;
3. instituições científicas e profissionais;
4. representantes de usuários e comunidades afetadas;
5. órgãos de controle e contratação pública;
6. provedores e trabalhadores da cadeia;
7. especialistas independentes em direitos humanos e acessibilidade.

O objetivo não é obter unanimidade para cada inferência. É definir padrões, pisos, métricas, deveres de informação e procedimentos de revisão.

10.3 Computação pública orientada ao interesse coletivo

Capacidade pública pode reduzir dependência e permitir pesquisa fora das prioridades comerciais. Deve ser concebida como infraestrutura de missão pública, não apenas como subsídio para empresas utilizarem recursos estatais sem contrapartida.

Princípios mínimos incluem:

- finalidade pública definida;
- acesso transparente e não discriminatório;
- prioridade baseada em critérios contestáveis, nunca em valor social da pessoa;
- cotas ou reservas para universidades, serviços públicos, pequenas organizações e problemas negligenciados;
- padrões ambientais de ciclo de vida;
- publicação de resultados e aprendizado, respeitados sigilos legítimos;
- interoperabilidade e uso de componentes abertos quando viável;
- prevenção de captura por poucos laboratórios;
- conselho independente e prestação de contas.

A computação pública não elimina provedores privados. Cria alternativa e referência. Sua existência melhora poder de negociação, capacidade de auditoria e continuidade.

10.4 Infraestrutura comum e pluralidade de modelos

Um ecossistema resiliente combina modelos locais, abertos, especializados e proprietários. Pesos abertos reduzem uma barreira, mas não entregam chips, energia, equipes, segurança, dados ou capacidade de retreinamento. A abertura deve ser avaliada em toda a cadeia.

Pluralidade reduz o risco de que uma única ontologia, política de conteúdo ou arquitetura de segurança defina o campo cognitivo nacional. Para ser efetiva, exige padrões de interface, formatos portáteis, documentação, conjuntos públicos de avaliação e equipes capazes de operar alternativas.

10.5 Direito de saber, contestar e escalar

Em sistemas que afetam direitos, o usuário deve poder:

- saber que classe de capacidade participou da decisão;
- conhecer limitações materialmente relevantes;
- solicitar revisão quando a capacidade puder ter sido insuficiente;
- obter canal humano em hipóteses definidas;

- contestar classificação ou roteamento;
- pedir preservação dos registros necessários;
- receber resposta sem retaliação comercial.

Esse **direito de escalar** não significa acesso automático e gratuito ao modelo mais amplo para qualquer solicitação. Significa procedimento para demonstrar que o piso foi insuficiente diante do risco, dos direitos ou da finalidade.

10.6 Dever de desacelerar e dever de ampliar

Governança de suficiência não é apenas restritiva. Em alguns casos, a autoridade deve impedir expansão injustificada; em outros, deve ampliar capacidade para corrigir desigualdade, garantir serviço ou enfrentar risco. Há, portanto, deveres simétricos:

- **dever de desacelerar**, quando o ganho marginal não justifica custo e dependência;
- **dever de ampliar**, quando economia ou escassez reduz proteção abaixo do piso devido.

A justiça do limiar depende dessa simetria. Sustentabilidade sem igualdade pode virar austeridade tecnológica; igualdade sem limites materiais pode reproduzir expansão insustentável.

10.7 Mapa de responsabilidades

Responsabilidade deve acompanhar controle e informação. Nenhum agente responde sozinho por toda a cadeia, mas nenhum pode ocultar-se atrás da complexidade coletiva.

Quadro 6 — Distribuição indicativa de responsabilidades

Agente	Controle predominante	Deveres centrais no TLSA
Fabricante ou desenvolvedor	arquitetura, treinamento e limites conhecidos	documentar desempenho, riscos, recursos e mudanças materiais
Provedor de nuvem ou infraestrutura	computação, disponibilidade, localização e consumo operacional	fornecer métricas verificáveis, continuidade, segurança e portabilidade
Integrador	seleção, configuração, dados e conexão com processos	demonstrar adequação, testar alternativas e preservar registros
Instituição usuária	finalidade, contexto e decisão organizacional	definir requisitos, avaliar impacto, supervisionar e garantir contestação
Poder público regulador ou contratante	normas, incentivos, licenças e compras	fixar pisos, fiscalizar externalidades, proteger concorrência e prover recurso
Auditor independente	verificação técnica e institucional	testar alegações, declarar limitações e comunicar conflito de interesses

Fonte: elaboração do autor.

Contratos podem distribuir tarefas, mas não extinguir dever legal. Quem controla informação essencial deve produzi-la; quem define a finalidade deve justificar o uso; quem possui poder de interromper ou corrigir deve agir diante de risco.

11 O BRASIL DIANTE DO LIMIAR: DESENVOLVIMENTO SEM SUBSTITUIÇÃO DE DEPENDÊNCIAS

11.1 Diagnóstico estrutural e dimensões de capacidade

O Brasil reúne vantagens relevantes: matriz elétrica comparativamente renovável, dimensão continental, mercado interno, universidades e instituições de pesquisa, sistemas públicos de saúde e justiça, diversidade linguística e cultural e tradição diplomática. Enfrenta, simultaneamente, dependência de semicondutores avançados, concentração de nuvem, desigualdade regional de infraestrutura, baixa transparência de fornecedores e descontinuidade de investimento científico.

A resposta não consiste em escolher abstratamente entre ecossistemas norte-americanos, chineses ou europeus. Parcerias múltiplas aumentam margem de negociação, mas soberania não resulta do número de fornecedores. Se todos exigirem componentes inacessíveis, padrões fechados e contratos sem portabilidade, a diversificação será apenas geográfica.

A política brasileira precisa distinguir cinco capacidades verificáveis:

1. **usar** — contratar e operar sistemas;
2. **avaliar** — testar qualidade, risco, impacto e custo total;
3. **auditar** — reconstruir comportamento e verificar obrigações;
4. **adaptar** — modificar modelos, dados e ferramentas para contextos nacionais;
5. **substituir** — manter a função quando fornecedor, componente ou condição geopolítica mudar.

O país pode avançar rapidamente na primeira e continuar dependente nas demais. O indicador mais exigente de soberania não é quantos sistemas foram adquiridos, mas quanto tempo e conhecimento seriam necessários para substituí-los sem paralisar o serviço.

11.2 Infraestrutura e políticas em curso

O Plano Brasileiro de Inteligência Artificial (PBIA), lançado em 2024, prevê cerca de R\$ 23 bilhões em quatro anos e inclui infraestrutura, formação, serviços públicos, inovação e governança. Documentos governamentais associam o plano à redução de dependência e à disponibilização compartilhada de supercomputação para universidades, centros de pesquisa e

empresas emergentes (MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO, 2025a).

O Brasil já possui ativos que impedem um diagnóstico de ausência total. O supercomputador Santos Dumont, instalado no Laboratório Nacional de Computação Científica, recebeu expansão de aproximadamente 20 *petaflops* e 68.064 núcleos e é apresentado pelo LNCC como o maior da América Latina dedicado à pesquisa científica. Em 2025, o MCTI também atualizou regras de ingresso no Sistema Nacional de Processamento de Alto Desempenho (SINAPAD) e articulou Centros Nacionais de Processamento de Alto Desempenho (LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA, 2025; MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO, 2025b).

Na cadeia industrial, a Lei nº 14.968/2024 instituiu o Programa Brasil Semicondutores, ampliando instrumentos do PADIS, inclusive para projeto de circuitos e serviços intensivos em conhecimento. A política é relevante, mas não equivale, por si, a domínio nacional das etapas mais avançadas de fabricação. É preciso medir adicionalidade tecnológica, conteúdo nacional, formação, propriedade intelectual, escala e resiliência, e não apenas habilitações e dispêndio fiscal (BRASIL, 2024; MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS, 2026).

O quadro revela base institucional real, porém fragmentada. PBIA, SINAPAD, LNCC, Serpro, Dataprev, universidades, política industrial e compras públicas não formam automaticamente um sistema coordenado de capacidade. Faltam inventário comum, métricas comparáveis, regra pública de prioridade, avaliação de concentração e plano de continuidade entre infraestruturas.

11.3 Data centers, energia e o risco de socialização do custo

A expansão de centros de dados tornou-se política econômica e energética. A Medida Provisória nº 1.318/2025, que instituiu o Redata, perdeu eficácia em fevereiro de 2026. O PL nº 278/2026, aprovado pela Câmara e ainda aguardando apreciação do Senado em 24 de julho de 2026, busca restabelecer o regime tributário. Em paralelo, o Ministério das Comunicações anunciou a preparação de políticas nacionais de data centers, cabos submarinos e inclusão digital, com o propósito de distribuir infraestrutura e reforçar resiliência (CONGRESSO NACIONAL, 2026; MINISTÉRIO DAS COMUNICAÇÕES, 2026).

A dimensão material não é marginal. Segundo a Empresa de Pesquisa Energética, os projetos de data centers com pedidos em processo no Ministério de Minas e Energia passaram de 19,8 GW, em setembro de 2025, para 26,2 GW, em novembro do mesmo ano. A própria

EPE adverte que esse montante não equivale a projetos efetivamente implantados e depende de transmissão, conectividade, financiamento e garantias de conexão. Ainda assim, a ordem de grandeza demonstra que política de IA também é planejamento elétrico e territorial (EMPRESA DE PESQUISA ENERGÉTICA, 2026).

Atrair data centers sem encadeamento pode transferir custos de rede, água e território e produzir pouco conhecimento substituível. Incentivos, licenciamento e conexão devem ser avaliados com contrapartidas proporcionais:

- transparência verificável sobre água, energia, emissões e equipamentos;
- adicionalidade energética e contribuição para expansão da rede;
- proteção de comunidades, bacias e usuários expostos a custo de oportunidade;
- circularidade e destino de equipamentos;
- formação, pesquisa e engenharia no país;
- acesso público ou científico a parcela da capacidade quando houver subsídio relevante;
- segurança, interoperabilidade, resposta a incidentes e plano de descontinuação;
- vedação à socialização automática de custos que permanecem privados nos benefícios.

“Energia renovável” é requisito importante, não conclusão. Fontes de baixo carbono também ocupam território, exigem transmissão e possuem usos concorrentes. O TLISA pergunta se o incremento é adicional, quem paga pela infraestrutura e qual função social recebe prioridade.

11.4 Regulação em construção e governança setorial

O PL nº 2.338/2023, aprovado pelo Senado, permanece em exame na Comissão Especial da Câmara. Sua tramitação demonstra que o marco geral brasileiro ainda não está concluído; por isso, o artigo não pode atribuir ao projeto força normativa vigente (CÂMARA DOS DEPUTADOS, 2025).

Enquanto isso, regimes setoriais avançam. A Resolução CNJ nº 615/2025 instituiu estrutura detalhada para IA no Poder Judiciário. A existência desse ato reduz o risco de vazio absoluto, mas também evidencia fragmentação: diferentes setores possuem capacidades técnicas e regras desiguais. Uma governança nacional do limiar deve interoperar com autoridades ambientais, energéticas, concorrenciais, de proteção de dados, de controle e setoriais, sem concentrar todas as decisões em um novo órgão tecnocrático.

O problema central não é ausência de qualquer política. É ausência de ligação operacional entre política industrial, energia, impacto territorial, capacidade pública,

contratação, direitos e substituição. O TLSA pretende servir como linguagem comum entre essas arenas, sem apagar suas competências.

11.5 Capacidade pública prioritária sem tecnocracia distributiva

Recursos públicos de computação são limitados. A alocação requer prioridades, mas não pode tornar-se concurso de prestígio institucional. Critérios defensáveis incluem adicionalidade pública, inexistência de alternativa acessível, risco social, abertura dos resultados, formação de capacidade local, diversidade regional e impacto material.

Universidades consolidadas possuem projetos legítimos; instituições regionais, serviços municipais, organizações sociais e pesquisadores independentes também precisam de vias de acesso. Reservas, rodízios, chamadas temáticas e apoio à formulação corrigem a vantagem de quem já domina editais e possui equipes especializadas.

O juízo de mérito deve recair sobre a relação entre projeto, evidência e interesse público, nunca sobre o valor da pessoa. Decisões devem ser motivadas, registradas e recorríveis. A governança precisa publicar dados agregados de concessões e recusas, utilização, filas e concentração. Sem isso, “excelência” pode apenas reproduzir acesso pretérito.

11.6 Compras estatais como política de soberania

O Estado é grande comprador e formador de mercado. Se cada órgão adquirir isoladamente solução fechada, aumentará o poder dos fornecedores. Catálogos comuns, cláusulas padronizadas, avaliações compartilhadas, ambientes de teste e arquitetura modular podem reduzir assimetria.

Compras públicas devem remunerar não somente acesso, mas **transferência de capacidade institucional**: documentação, treinamento, portabilidade, integração, registros, segurança e saída. A instituição precisa conservar conhecimento suficiente para fiscalizar e substituir. Caso contrário, terceiriza simultaneamente operação e compreensão.

Um cadastro interoperável de sistemas usados em funções públicas, com publicidade graduada, permitiria visualizar concentração, gastos, versões, dependências e incidentes. Não seria autorização universal, mas instrumento de inteligência estatal e coordenação.

11.7 Agenda brasileira de soberania sobre o limiar

Uma política articulada pode incluir:

1. inventário nacional de capacidade computacional pública e contratada;
2. métricas comuns de energia, água, carbono, desempenho, utilização e dependência;

3. rede federada de computação para pesquisa e serviços públicos;
4. laboratórios independentes de avaliação e auditoria;
5. padrões de portabilidade, registros e interoperabilidade;
6. formação técnica distribuída regionalmente;
7. política de dados públicos com governança, qualidade e proteção;
8. exigências de ciclo de vida em licenciamento, incentivos e contratação;
9. observatório de concentração da cadeia de IA;
10. participação social na definição de prioridades e revisão de limiares;
11. testes periódicos de continuidade e substituição em serviços essenciais;
12. avaliação pública de contrapartidas quando energia, crédito, tributo ou infraestrutura forem subsidiados.

O objetivo não é autossuficiência integral. É impedir que qualquer fornecedor, país ou arquitetura se torne indispensável ao funcionamento do Estado e da economia.

12 OBJEÇÕES, LIMITES E RISCOS DA PROPOSTA

12.1 “O teste desacelerará a inovação”

Qualquer obrigação de justificar custos adiciona tempo. O efeito precisa ser calibrado por risco e escala. Não se propõe aplicar avaliação completa a cada consulta individual. O TLSA é mais intenso em decisões de infraestrutura, desenvolvimento de fronteira, contratação coletiva, serviços essenciais e sistemas de alto impacto.

Inovação sem análise também possui custo: contratos inúteis, projetos abandonados, dependência, acidentes e infraestrutura ociosa. Definição prévia de finalidade e requisitos pode acelerar soluções adequadas e abrir espaço para modelos menores, especializados e locais que a corrida maximalista marginaliza.

12.2 “Não é possível medir o impacto com precisão”

De fato, dados ambientais de IA são incompletos e metodologias variam. Comparações entre modelos podem ser frágeis, especialmente quando fornecedores não divulgam arquitetura, utilização e matriz energética. O artigo não pressupõe precisão inexistente.

A resposta jurídica à incerteza é padronizar métricas, declarar intervalos, registrar hipóteses e atribuir ônus informacional a quem controla os dados. Decisões podem usar

faixas, cenários e indicadores físicos separados. Esperar mensuração perfeita favorece quem lucra com a opacidade.

O teste também evita falso rigor: não converte todas as dimensões em pontuação única. Um quadro multicritério preserva conflitos e permite motivação pública.

12.3 “O maior modelo pode ser mais eficiente no resultado final”

É possível. Um modelo superior pode resolver tarefa em menos tentativas, reduzir trabalho humano, evitar erro ou substituir cadeia mais intensiva. Parâmetros ou energia por inferência isolada não determinam impacto total. Por isso, o TLSA compara resultados funcionais e ciclo de vida, não presume que menor seja sempre melhor.

A economia de aperfeiçoamento não é culto à miniaturização. Ela exige prova contextual. Em decisão clínica complexa, pesquisa científica ou resposta a desastre, capacidade adicional pode ser claramente necessária. A regra contém tanto o excesso quanto a insuficiência.

12.4 “O mercado ajustará preço e capacidade”

Preços sinalizam escassez, mas refletem poder de mercado, subsídios, externalidades e disposição a pagar. Grupos com maior renda podem consumir mais capacidade mesmo quando benefício social marginal é baixo; serviços essenciais podem receber menos porque seu orçamento é restrito. Energia, água e conhecimento público podem ser subsidiados sem aparecer integralmente no preço.

Além disso, concentração e custos de troca reduzem a disciplina concorrencial. A OCDE identifica barreiras de entrada, relações verticais e obstáculos à mudança na infraestrutura de IA (OECD, 2025). O mercado participa da alocação, mas não substitui direitos, política ambiental ou finalidade pública.

12.5 “O Estado pode usar a suficiência para censurar e controlar”

Esse é um dos riscos centrais. Uma autoridade poderia rotular perguntas como inúteis, restringir capacidade de opositores, vigiar finalidades ou monopolizar computação. A proposta só é defensável com salvaguardas:

- critérios dirigidos ao incremento e à tarefa, não à ideologia ou identidade;
- piso universal de acesso;
- motivação e recurso;
- pluralidade de provedores e instituições;
- fiscalização judicial e independente;

- transparência agregada;
- proteção de privacidade;
- proibição de discriminação política.

Computação pública deve ampliar alternativas, não extinguir privadas. Policentrismo e direito de saída reduzem captura estatal e empresarial.

12.6 “O roteamento eficiente ameaça privacidade e liberdade”

Classificar solicitações para escolher modelos pode exigir analisar conteúdo, perfil e contexto. Essa análise cria vigilância adicional. O sistema deve aplicar minimização, processamento local quando possível, separação de finalidades, retenção limitada e proibição de uso secundário não autorizado.

Crítérios podem ser baseados em requisitos técnicos fornecidos pelo próprio usuário ou pela aplicação, evitando inferências sobre valor social. Em alto impacto, o roteamento deve ser auditável. Economia de recursos não justifica expansão invisível da coleta.

12.7 “Capacidade aberta resolverá a concentração”

Pesos abertos são importantes, mas insuficientes. Treinar, ajustar, proteger e executar modelos exige computação, dados, equipes e energia. Licenças podem impor restrições; documentação pode faltar; cadeias de hardware e nuvem permanecem concentradas. Widder, Whittaker e West (2024) mostram que a abertura de sistemas de IA possui múltiplas camadas e frequentemente é parcial.

Abertura deve ser avaliada como propriedade do ecossistema: pesos, código, dados, avaliação, hardware, governança e possibilidade material de modificação. Soberania formal sem capacidade material é dependência com licença mais permissiva.

12.8 “O limiar será rapidamente obsoleto”

Será, se tratado como número fixo. O TLSA é procedimento e inclui revisão. Melhorias algorítmicas, novas ameaças, preços, matriz energética e experiência de uso mudam o suficiente. Contratos e normas devem prever gatilhos objetivos, versionamento e reavaliação.

Essa instabilidade não invalida o conceito. Direito já governa padrões dinâmicos por diligência, estado da técnica, proporcionalidade e gestão de risco. O desafio é conservar memória e evidência para que revisão seja aprendizado, não improvisação.

12.9 Limites desta investigação

Este artigo possui natureza teórico-conceitual. Não apresenta inventário completo da capacidade brasileira, cálculo de impacto ambiental de modelos nem validação empírica do TLSA. Também não resolve conflitos entre sigilo industrial e auditoria, competências federativas, tributação de externalidades ou desenho financeiro de infraestrutura pública.

As expressões propostas precisam ser confrontadas com casos. Há risco de sobreposição terminológica, burocratização e captura. A formulação foi deliberadamente construída como hipótese normativa aberta à refutação, não como sistema fechado.

13 AGENDA DE PESQUISA E VALIDAÇÃO FUTURA

A consolidação ou refutação das categorias propostas exige programa empírico e comparado. Oito frentes parecem prioritárias.

13.1 Cartografia da cadeia e dos pontos de indispensabilidade

É necessário mapear chips, nuvem, modelos, energia, redes, dados, equipes e contratos usados por setores estratégicos brasileiros. O objetivo não é apenas identificar fornecedores, mas dependências sem substituto, prazos de migração e conhecimento interno.

Perguntas de pesquisa:

- quais funções estatais já dependem de um único ecossistema;
- quantas camadas chamadas “nacionais” permanecem controladas externamente;
- quais componentes podem ser substituídos sem perda crítica;
- onde subsídios públicos financiam capacidade privada sem contrapartida.

13.2 Medição territorial de externalidades

Estudos nacionais devem relacionar centros de dados a bacias hidrográficas, rede elétrica, preço de energia, emprego, arrecadação, resíduos e uso do solo. Médias globais são inadequadas para conflitos locais.

Hipótese: a distribuição territorial dos custos difere da distribuição societária e fiscal dos benefícios, produzindo zonas de extração digital.

13.3 Auditoria de estratificação por interface

Experimentos podem comparar respostas, latência, ferramentas, contexto e segurança entre planos, regiões e perfis. O objetivo é identificar redução silenciosa, discriminação e assimetria de informação.

Hipótese: interfaces semelhantes ocultam diferenças materiais de capacidade que usuários não conseguem perceber ou contestar.

13.4 Validação do TLSA

O teste deve ser aplicado retrospectivamente e em pilotos de saúde, justiça, educação, administração e pesquisa. Avaliadores independentes podem examinar consistência, custo de aplicação, divergência entre equipes e impacto sobre escolhas.

Questões:

- quais critérios alteram decisões;
- quais evidências são inviáveis hoje;
- o teste favorece alternativas menores sem reduzir qualidade;
- quais itens funcionam como piso e quais admitem ponderação;
- como calibrar intensidade por risco e escala.

13.5 Direito comparado da infraestrutura de IA

A pesquisa deve comparar União Europeia, Estados Unidos, China, Índia, países latino-americanos e iniciativas multilaterais. Mais do que catalogar leis de modelos, precisa examinar energia, licenciamento, concorrência, compras públicas, exportação de chips e capacidade pública.

Hipótese: a governança internacional está fragmentada entre segurança, direitos, desenvolvimento e cadeia de suprimentos, enquanto a autoridade sobre o limiar permanece dispersa e frequentemente privada.

13.6 Economia política da agenda científica

Bibliometria e entrevistas podem investigar como acesso a computação altera temas, métodos, colaborações e publicação no Brasil. A divisão computacional não se mede apenas por horas de GPU, mas por problemas que deixam de ser pesquisados.

Hipótese: dependência de créditos e parcerias orienta a ciência a tarefas compatíveis com plataformas, reduzindo pluralidade e capacidade crítica.

13.7 Influência cognitiva estrutural de segunda ordem

A pesquisa anterior sobre influência cognitiva estrutural pode ser ampliada para efeitos institucionais temporais: como a dependência de IA modifica competência residual, memória organizacional, linguagem e possibilidade de saída.

Hipótese: a arquitetura que distribui capacidade no presente também reorganiza a capacidade humana e institucional de decidir no futuro.

13.8 Proposições orientadoras

Investigações futuras poderão testar as seguintes proposições:

1. eficiência sem suficiência tende a produzir rebote;
2. suficiência sem justiça tende a produzir austeridade e estratificação;
3. acesso sem auditabilidade tende a produzir dependência;
4. abertura sem infraestrutura tende a produzir soberania formal;
5. diversificação sem interoperabilidade tende a diversificar dependências;
6. computação pública sem governança tende a reproduzir captura;
7. escala sem demonstração marginal tende a converter possibilidade técnica em necessidade política;
8. soberania exige poder sobre o limiar, e não apenas presença em fóruns ou acesso contratual.

14 CONSIDERAÇÕES FINAIS

A inteligência artificial tornou-se uma infraestrutura de produção de linguagem, conhecimento, previsão e decisão. Sua expansão depende de recursos materiais finitos e de cadeias concentradas. A questão ambiental não pode ser separada da questão distributiva; a questão distributiva não pode ser separada da autoridade; e a autoridade não pode ser separada da autonomia.

Green AI mostrou que desempenho precisa ser acompanhado de custo. IA frugal demonstrou que restrição pode estimular soluções eficientes e acessíveis. A suficiência digital deslocou o olhar do custo unitário para o impacto absoluto. O *right-sizing* mostrou que o maior modelo nem sempre é necessário. A literatura sobre divisão computacional revelou que acesso à infraestrutura organiza quem pode pesquisar e definir a fronteira. As propostas de computação pública indicaram alternativas ao monopólio privado.

O presente artigo procurou integrar esses avanços e tornar visível uma decisão anterior: quem define o suficiente. Para isso, empregou cinco categorias. A economia de aperfeiçoamento examina a justificativa do incremento marginal. A economia política da capacidade algorítmica mapeia poder, propriedade, destinação e externalidades. A privatização do limiar de suficiência identifica a transferência silenciosa da autoridade para fornecedores. A suficiência algorítmica distributiva combina teto de desperdício e piso de

direitos. A soberania sobre o limiar exige capacidade normativa, material, tecnológica e decisória para definir, revisar e abandonar ecossistemas.

O Teste do Limiar de Suficiência Algorítmica traduz essas categorias em procedimento e especializa a proporcionalidade para a materialidade, o piso distributivo, a dependência e a revisão dinâmica. Nesta versão, permanece estrutura heurística ainda não validada. O caso-tipo demonstrou como ele pode alterar a instrução de uma contratação, sem provar consistência entre avaliadores ou superioridade decisória. Essa validação integra a agenda de pesquisa, e não pode ser presumida pela força retórica da fórmula:

Nem mais capacidade do que a finalidade justifica; nem menos capacidade do que o risco e os direitos exigem.

Não se trata de reservar inteligência a uma elite nem de vigiar o valor das perguntas humanas. Pessoas não são unidades de otimização. O objeto do limite é o incremento injustificado, a externalização de custos e a autoridade unilateral. O objeto da garantia é a capacidade adequada ao exercício de direitos, à segurança, à ciência e à finalidade pública.

No plano jurídico, a proposta também conserva limites. Finalidade, necessidade, planejamento, ciclo de vida, proteção de dados, prevenção, transparência, concorrência e governança setorial já possuem fundamentos positivos. Um dever geral de revelar classes de capacidade, assegurar escalonamento, impor interoperabilidade ou distribuir computação exige norma competente. O TLISA pode instruir decisões e contratos hoje; sua coerção futura depende de incorporação válida, competência e devido processo.

Uma sociedade que apenas compra respostas não controla a infraestrutura que forma suas perguntas. Um Estado que utiliza modelos sem capacidade de avaliar, adaptar e substituir preserva a assinatura formal da decisão, mas perde parte de seu domínio material. E um país que fornece energia, água, dados e mercado, enquanto importa critérios de finalidade e acesso, participa da cadeia sem governar seu lugar nela.

A escassez não determinará sozinha o futuro da IA. Instituições decidirão como ela será administrada. Se essas decisões permanecerem opacas e concentradas, o limite necessário será privatizado. Se forem submetidas a direitos, evidência, contestação e pluralidade, a finitude poderá deixar de ser instrumento de exclusão e tornar-se princípio de responsabilidade.

Em última análise, soberania em inteligência artificial não se mede pelo acesso ocasional ao sistema mais poderoso. Mede-se pela capacidade coletiva de decidir quando o poder adicional é necessário, de impedir que seus custos sejam impostos aos ausentes e de assegurar que ninguém receba menos do que a dignidade, o risco e o direito exigem.

Referências

- ACEMOGLU, Daron; JOHNSON, Simon. *Power and progress: our thousand-year struggle over technology and prosperity*. New York: PublicAffairs, 2023.
- AHMED, Nur; WAHED, Muntasir. The de-democratization of AI: deep learning and the compute divide in artificial intelligence research. *arXiv*, 2020. DOI: 10.48550/arXiv.2010.15581. Disponível em: <https://arxiv.org/abs/2010.15581>. Acesso em: 24 jul. 2026.
- AHMED, Nur; WAHED, Muntasir; THOMPSON, Neil C. The growing influence of industry in AI research. *Science*, Washington, v. 379, n. 6635, p. 884-886, 2023. DOI: 10.1126/science.ade2420.
- ALEX, Robert. On balancing and subsumption: a structural comparison. *Ratio Juris*, v. 16, n. 4, p. 433-449, 2003. DOI: 10.1111/j.0952-1917.2003.00244.x.
- ARGA, Ludovic et al. Frugal AI: introduction, concepts, development and open questions. *ACM SIGKDD Explorations Newsletter*, New York, v. 27, n. 1, p. 72-111, 2025. DOI: 10.1145/3748239.3748247.
- ÁVILA, Humberto. *Teoria dos princípios: da definição à aplicação dos princípios jurídicos*. 20. ed. São Paulo: Malheiros, 2021.
- AXELSEN, David V.; NIELSEN, Lasse. Essentially enough: elements of a plausible account of sufficientarianism. In: FOURIE, Carina; RID, Annette (ed.). *What is enough? Sufficiency, justice, and health*. Oxford: Oxford University Press, 2016. p. 101-118. DOI: 10.1093/acprof:oso/9780199385263.003.0007.
- AXELSEN, David V.; SHIELDS, Liam. How thresholds matter: on the bounds and demands of justice. *Utilitas*, v. 37, n. 3, p. 225-245, 2025. DOI: 10.1017/S095382082500010X.
- BARAK, Aharon. *Proportionality: constitutional rights and their limitations*. Cambridge: Cambridge University Press, 2012.
- BAUMANN, Joachim; HANNÁK, Anikó; HEITZ, Christoph. Enforcing group fairness in algorithmic decision making: utility maximization under sufficiency. In: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2022, Seoul. *Proceedings [...]*. New York: ACM, 2022. DOI: 10.1145/3531146.3534645.
- BRASIL. [Constituição (1988)]. *Constituição da República Federativa do Brasil de 1988*. Brasília, DF: Presidência da República, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 24 jul. 2026.
- BRASIL. Emenda Constitucional nº 115, de 10 de fevereiro de 2022. Altera a Constituição Federal para incluir a proteção de dados pessoais entre os direitos e garantias fundamentais. *Diário Oficial da União*: Brasília, DF, 11 fev. 2022. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc115.htm. Acesso em: 24 jul. 2026.
- BRASIL. Lei nº 6.938, de 31 de agosto de 1981. Dispõe sobre a Política Nacional do Meio Ambiente. *Diário Oficial da União*: Brasília, DF, 2 set. 1981. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/16938.htm. Acesso em: 24 jul. 2026.
- BRASIL. Lei nº 8.078, de 11 de setembro de 1990. Dispõe sobre a proteção do consumidor e dá outras providências. *Diário Oficial da União*: Brasília, DF, 12 set. 1990. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/18078compilado.htm. Acesso em: 24 jul. 2026.
- BRASIL. Lei nº 12.529, de 30 de novembro de 2011. Estrutura o Sistema Brasileiro de Defesa da Concorrência e dispõe sobre a prevenção e repressão às infrações contra a ordem econômica. *Diário Oficial da União*: Brasília, DF, 1 dez. 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/ato2011-2014/2011/lei/112529.htm. Acesso em: 24 jul. 2026.
- BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). *Diário Oficial da União*: Brasília, DF, 15 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/113709.htm. Acesso em: 24 jul. 2026.
- BRASIL. Lei nº 14.133, de 1º de abril de 2021. Lei de Licitações e Contratos Administrativos. *Diário Oficial da União*: Brasília, DF, 1 abr. 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/ato2019-2022/2021/lei/114133.htm. Acesso em: 24 jul. 2026.

- BRASIL. Lei nº 14.968, de 11 de setembro de 2024. Aperfeiçoa a política industrial para os setores de tecnologias da informação e comunicação e de semicondutores e cria o Programa Brasil Semicondutores. *Diário Oficial da União*: Brasília, DF, 12 set. 2024. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2024/lei/114968.htm. Acesso em: 24 jul. 2026.
- CÂMARA DOS DEPUTADOS. *Projeto de Lei nº 2.338, de 2023: ficha de tramitação*. Brasília, DF: Câmara dos Deputados, 2025. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2520999>. Acesso em: 24 jul. 2026.
- CHEN, Lingjiao; ZAHARIA, Matei; ZOU, James. FrugalGPT: how to use large language models while reducing cost and improving performance. *arXiv*, 2023. DOI: 10.48550/arXiv.2305.05176. Disponível em: <https://arxiv.org/abs/2305.05176>. Acesso em: 24 jul. 2026.
- CHIEN, Andrew A. et al. Strategies and design for increasing AI sustainability. *Nature Reviews Clean Technology*, 2026. DOI: 10.1038/s44359-026-00195-w.
- CONGRESSO NACIONAL. *Projeto de Lei nº 278, de 2026: Regime Especial de Tributação para Serviços de Datacenter — Redata*. Brasília, DF: Congresso Nacional, 2026. Disponível em: <https://www.congressonacional.leg.br/materias/materias-bicamerais/-/ver/pl-278-2026>. Acesso em: 24 jul. 2026.
- CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 615, de 11 de março de 2025. Estabelece diretrizes para o desenvolvimento, utilização e governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário. *DJe/CNJ*, n. 54, p. 2-17, 14 mar. 2025. Texto compilado. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/6001>. Acesso em: 24 jul. 2026.
- CRAWFORD, Kate. *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press, 2021.
- DAVIES, Matt; VIPRA, Jai. *Computing commons: designing public compute for people and society*. London: Ada Lovelace Institute, 2025. Disponível em: <https://www.adalovelaceinstitute.org/report/computing-commons/>. Acesso em: 24 jul. 2026.
- DICICCIO, Cyrus et al. Detection and mitigation of algorithmic bias via predictive rate parity. In: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2023, Chicago. *Proceedings [...]*. New York: ACM, 2023. DOI: 10.1145/3593013.3594117.
- DING, Dujian et al. Hybrid LLM: cost-efficient and quality-aware query routing. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, 2024. *Proceedings [...]*. [S. l.]: ICLR, 2024. Disponível em: <https://www.microsoft.com/en-us/research/publication/hybrid-llm-cost-efficient-and-quality-aware-query-routing/>. Acesso em: 24 jul. 2026.
- EMPRESA DE PESQUISA ENERGÉTICA. REDATA: EPE acelera planejamento da rede para suportar crescimento recorde de data centers. Rio de Janeiro: EPE, 2026. Disponível em: <https://www.epe.gov.br/pt/imprensa/noticias/redata-epe-acelera-planejamento-da-rede-para-suportar-crescimento-recorde-de-data-centers>. Acesso em: 24 jul. 2026.
- FEENBERG, Andrew. *Questioning technology*. London: Routledge, 1999.
- FOURIE, Carina; RID, Annette (ed.). *What is enough? Sufficiency, justice, and health*. Oxford: Oxford University Press, 2016. DOI: 10.1093/acprof:oso/9780199385263.001.0001.
- FRANKFURT, Harry G. Equality as a moral ideal. *Ethics*, v. 98, n. 1, p. 21-43, 1987.
- GENOVA, Jandislei Antonio. *Autonomia decisória na era da IA: formação da vontade, influência cognitiva estrutural e responsabilidade jurídica em ambientes algorítmicos*. São Paulo: Editora Dialética, 2026. ISBN 978-65-288-0467-2.
- HAASE, Jennifer; KLESSASCHECK, Finn; MENDLING, Jan; POKUTTA, Sebastian. Sustainability via LLM right-sizing. *arXiv*, 2025. DOI: 10.48550/arXiv.2504.13217. Disponível em: <https://arxiv.org/abs/2504.13217>. Acesso em: 24 jul. 2026.
- HENDERSON, Peter et al. Towards the systematic reporting of the energy and carbon footprints of machine learning. *Journal of Machine Learning Research*, v. 21, n. 248, p. 1-43, 2020.
- IEA — INTERNATIONAL ENERGY AGENCY. *Energy and AI*. Paris: IEA, 2025. Disponível em: <https://www.iea.org/reports/energy-and-ai>. Acesso em: 24 jul. 2026.

IEA — INTERNATIONAL ENERGY AGENCY. *Key questions on energy and AI*. Paris: IEA, 2026. Disponível em: <https://www.iea.org/reports/key-questions-on-energy-and-ai>. Acesso em: 24 jul. 2026.

LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA. Pesquisa desenvolvida com apoio do Supercomputador Santos Dumont é apresentada na Conferência IPDPS25. Petrópolis: LNCC, 2025. Disponível em: <https://www.gov.br/lncc/pt-br/assuntos/noticias/ultimas-noticias-1/pesquisa-desenvolvida-com-apoio-do-supercomputador-santos-dumont-e-apresentada-na-conferencia-ipdps25>. Acesso em: 24 jul. 2026.

LEHDONVIRTA, Vili. *Cloud empires: how digital platforms are overtaking the state and how we can regain control*. Cambridge, MA: MIT Press, 2022.

LEHDONVIRTA, Vili; WÚ, Bóxi; HAWKINS, Zoe. Compute North vs. Compute South: the uneven possibilities of compute-based AI governance around the globe. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, v. 7, n. 1, p. 828-838, 2024. DOI: 10.1609/aies.v7i1.31683.

LUCCIONI, Alexandra Sasha; JERNITE, Yacine; STRUBELL, Emma. Power hungry processing: watts driving the cost of AI deployment? In: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2024, Rio de Janeiro. *Proceedings [...]*. New York: ACM, 2024. p. 85-99. DOI: 10.1145/3630106.3658542.

LUITSE, Dieuwertje; DENKENA, Wiebke. The great Transformer: examining the role of large language models in the political economy of AI. *Big Data & Society*, v. 8, n. 2, 2021. DOI: 10.1177/20539517211047734.

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO. PBIA projeta Brasil como referência em inteligência artificial responsável. Brasília, DF: MCTI, 2025a. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2025/08/pbia-projeta-brasil-como-referencia-em-inteligencia-artificial-responsavel>. Acesso em: 24 jul. 2026.

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO. Entre chips, dados e inteligência artificial, 2025 marcou o avanço da soberania tecnológica do Brasil. Brasília, DF: MCTI, 2025b. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2025/12/entre-chips-dados-e-inteligencia-artificial-2025-marcou-o-avanco-da-soberania-tecnologica-do-brasil>. Acesso em: 24 jul. 2026.

MINISTÉRIO DAS COMUNICAÇÕES. Ministério das Comunicações prepara políticas nacionais de infraestrutura para tornar Brasil referência em inclusão digital. Brasília, DF: MCom, 2026. Disponível em: <https://www.gov.br/mcom/pt-br/noticias/2026/maio/ministerio-das-comunicacoes-prepara-politicas-nacionais-de-infraestrutura-para-tornar-brasil-referencia-em-inclusao-digital>. Acesso em: 24 jul. 2026.

MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS. Programa Brasil Semicondutores (Brasil Semicon). Brasília, DF: MDIC, 2026. Disponível em: <https://www.gov.br/mdic/pt-br/composicao/sdic/dial/complexo-eletroeletronico-e-de-semicondutores/programa-brasil-semicondutores-brasil-semicon>. Acesso em: 24 jul. 2026.

OECD — ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *A blueprint for building national compute capacity for artificial intelligence*. Paris: OECD Publishing, 2023. DOI: 10.1787/876367e3-en.

OECD — ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. Competition in artificial intelligence infrastructure. *OECD Roundtables on Competition Policy Papers*, Paris, n. 330, 2025. DOI: 10.1787/623d1874-en.

ONG, Isaac et al. RouteLLM: learning to route LLMs with preference data. *arXiv*, 2024. DOI: 10.48550/arXiv.2406.18665. Disponível em: <https://arxiv.org/abs/2406.18665>. Acesso em: 24 jul. 2026.

OSTROM, Elinor. *Governing the commons: the evolution of institutions for collective action*. Cambridge: Cambridge University Press, 1990.

PARASURAMAN, Raja; MANZEY, Dietrich H. Complacency and bias in human use of automation: an attentional integration. *Human Factors*, v. 52, n. 3, p. 381-410, 2010. DOI: 10.1177/0018720810376055.

RICHARDSON, Katherine et al. Earth beyond six of nine planetary boundaries. *Science Advances*, Washington, v. 9, n. 37, eadh2458, 2023. DOI: 10.1126/sciadv.adh2458.

ROCKSTRÖM, Johan et al. A safe operating space for humanity. *Nature*, London, v. 461, p. 472-475, 2009. DOI: 10.1038/461472a.

- SANTARIUS, Tilman et al. Digital sufficiency: conceptual considerations for ICTs on a finite planet. *Annals of Telecommunications*, Paris, v. 78, p. 277-295, 2023. DOI: 10.1007/s12243-022-00914-x.
- SASTRY, Girish et al. Computing power and the governance of artificial intelligence. *arXiv*, 2024. DOI: 10.48550/arXiv.2402.08797. Disponível em: <https://arxiv.org/abs/2402.08797>. Acesso em: 24 jul. 2026.
- SCHWARTZ, Roy; DODGE, Jesse; SMITH, Noah A.; ETZIONI, Oren. Green AI. *Communications of the ACM*, New York, v. 63, n. 12, p. 54-63, 2020. DOI: 10.1145/3381831.
- STRUBELL, Emma; GANESH, Ananya; MCCALLUM, Andrew. Energy and policy considerations for deep learning in NLP. In: ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 57., 2019, Florence. *Proceedings [...]*. Florence: Association for Computational Linguistics, 2019. p. 3645-3650. DOI: 10.18653/v1/P19-1355.
- SUPREMO TRIBUNAL FEDERAL. Referendo em medida cautelar nas ações diretas de inconstitucionalidade nº 6.387, 6.388, 6.389, 6.390 e 6.393. Relatora: Ministra Rosa Weber. Brasília, DF, 7 maio 2020. *Diário da Justiça Eletrônico*, 2 jun. 2020. Disponível em: https://portal.stf.jus.br/textos/verTexto.asp?pagina=resumocovid_adi&servico=resumocovid. Acesso em: 24 jul. 2026.
- THALER, Richard H.; SUNSTEIN, Cass R. *Nudge: improving decisions about health, wealth, and happiness*. New Haven: Yale University Press, 2008.
- TRIBUNAL DE CONTAS DA UNIÃO. Acórdão nº 157/2024 — Plenário. Relator: Ministro Antonio Anastasia. Processo TC 010.613/2023-4. Sessão de 7 fev. 2024. Brasília, DF: TCU, 2024. Disponível em: <https://pesquisa.apps.tcu.gov.br/documento/acordao-completo/%2A/NUMACORDAO%253A157%2520ANOACORDAO%253A2024%2520/DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0>. Acesso em: 24 jul. 2026.
- UNEP — UNITED NATIONS ENVIRONMENT PROGRAMME. *Artificial intelligence (AI) end-to-end: the environmental impact of the full AI lifecycle needs to be comprehensively assessed*. Nairobi: UNEP, 2024. Disponível em: <https://www.unep.org/resources/report/artificial-intelligence-ai-end-end-environmental-impact-full-ai-lifecycle-needs-be>. Acesso em: 24 jul. 2026.
- UNESCO — UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION. *Recommendation on the ethics of artificial intelligence*. Paris: UNESCO, 2021. Disponível em: <https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence>. Acesso em: 24 jul. 2026.
- UNIÃO EUROPEIA. Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que cria regras harmonizadas em matéria de inteligência artificial. *Jornal Oficial da União Europeia*, 12 jul. 2024. Disponível em: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>. Acesso em: 24 jul. 2026.
- VIPRA, Jai; MYERS WEST, Sarah. *Computational power and AI*. New York: AI Now Institute, 2023. Disponível em: <https://ainowinstitute.org/publications/compute-and-ai>. Acesso em: 24 jul. 2026.
- WIDDER, David Gray; WHITTAKER, Meredith; MYERS WEST, Sarah. Why “open” AI systems are actually closed, and why this matters. *Nature*, London, v. 635, p. 827-833, 2024. DOI: 10.1038/s41586-024-08141-1.
- WINNER, Langdon. Do artifacts have politics? *Daedalus*, Cambridge, MA, v. 109, n. 1, p. 121-136, 1980.
- YEUNG, Karen. “Hypernudge”: big data as a mode of regulation by design. *Information, Communication & Society*, v. 20, n. 1, p. 118-136, 2017. DOI: 10.1080/1369118X.2016.1186713.